

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 12:44:35 ; Search time 58.24 Seconds
(without alignments)
3080.146 Million cell updates/sec

Title: US-08-940-544-3
 Perfect score: 717
 Sequence: 1 cagctgaaagctgcagcagctc.....cacagttcggaaatgaaacag 717

Scoring table: IDENTITY NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:★

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query % | | | DB | ID | Description |
|------------|---------|-------|--------|----|--------|---------------------|
| | Score | Match | Length | | | |
| 1 | 715 | 99.7 | 717 | 1 | T8G309 | Single chain anti- |
| 2 | 703 | 98.0 | 1175 | 1 | T8G311 | Single chain anti- |
| 3 | 523.8 | 73.1 | 894 | 1 | X33931 | HBV specific single |
| 4 | 521.6 | 72.7 | 726 | 1 | T04023 | Anti-EGFR single c |
| 5 | 520.6 | 72.6 | 726 | 1 | T04025 | Anti-EGFR single c |
| 6 | 520.4 | 72.6 | 732 | 1 | V23579 | Firefly luciferase |
| 7 | 517.8 | 72.2 | 726 | 1 | V23580 | Firefly luciferase |
| 8 | 517.4 | 72.2 | 2364 | 1 | V23579 | Firefly luciferase |
| 9 | 512.6 | 71.5 | 725 | 1 | V00687 | Fusion gene sequen |
| 10 | 507.2 | 70.7 | 810 | 1 | T04024 | Anti-EGFR single c |
| 11 | 504.6 | 70.4 | 732 | 1 | Q90663 | MFE-23 antibody co |
| 12 | 504.6 | 70.4 | 732 | 1 | T04020 | Anti-EGFR single c |
| 13 | 503 | 70.2 | 732 | 1 | T04021 | Anti-EGFR single c |
| 14 | 501.6 | 70.0 | 1679 | 1 | T58130 | Single chain anti- |
| 15 | 501.6 | 70.0 | 1679 | 1 | V08176 | H22-anti-CEA anti- |
| 16 | 499.8 | 69.7 | 1797 | 1 | X00812 | Vm-V1-linker-PE40 |
| 17 | 487 | 67.9 | 738 | 1 | T04022 | Anti-EGFR single c |
| 18 | 471 | 65.7 | 843 | 1 | V72069 | Fusion protein p1c |
| 19 | 471 | 65.7 | 1998 | 1 | V72075 | Fusion protein p1c |
| 20 | 460.6 | 64.2 | 1794 | 1 | V73337 | Fusion protein p1c |
| 21 | 452.6 | 63.1 | 843 | 1 | X01657 | Mouse bispecific a |
| 22 | 432.8 | 60.4 | 6115 | 1 | T62335 | Anti-B7.1/B7.2 bis |
| 23 | 425.4 | 59.3 | 828 | 1 | X01661 | Eukaryotic express |
| 24 | 410.8 | 57.3 | 738 | 1 | Q80468 | Anti-B7.1/anti-b7. |
| 25 | 410.8 | 57.3 | 729 | 1 | X01378 | Single chain antib |
| 26 | 410 | 57.2 | 1807 | 1 | V80290 | Murine anti-574 an |
| 27 | 407 | 56.8 | 1467 | 1 | V80291 | Anti-574 single ch |
| 28 | 406.8 | 56.7 | 1518 | 1 | V80292 | Human B7-1.574.1.1 |
| 29 | 406.8 | 56.7 | 2090 | 1 | V80294 | B7-1/scFv specific |
| 30 | 405.8 | 56.6 | 1725 | 1 | X01651 | ScFv-IgE1 fusion c |
| 31 | 400 | 55.8 | 759 | 1 | V10378 | Bispecific tetra |
| 32 | 397 | 55.4 | 711 | 1 | Q37461 | Anti-CD86 monoclon |
| 33 | 390.6 | 54.5 | 1314 | 1 | V09256 | VH N010/12.5-Vk NQ |
| 34 | 389.2 | 54.3 | 1668 | 1 | X01652 | Nucleotide sequenc |
| 35 | 387.6 | 54.1 | 744 | 1 | V11399 | Bispecific tetra |
| 36 | 386.8 | 53.9 | 708 | 1 | Q37459 | Human CD30 binding |
| 37 | 381 | 53.1 | 720 | 1 | V54790 | VH N02/12.4-Vk NQ1 |
| 38 | 371.2 | 51.8 | 1065 | 1 | T94971 | scFv comprising he |
| 39 | 370.8 | 51.7 | 2019 | 1 | V72059 | R. pipiens recombi |
| 40 | 370.8 | 51.7 | 2025 | 1 | V72064 | Plasmid pNG4/55.l |
| 41 | 370.8 | 51.7 | 864 | 1 | V72072 | Fusion protein p1c |
| 42 | 368.2 | 51.4 | 831 | 1 | Q62957 | Fusion protein p1c |
| 43 | 357 | 49.8 | 720 | 1 | V26770 | Anti-influenza N10 |

| | 44 | 348.8 | 48.6 | 2012 | 1 | Q28263 | Fv(FRP51)-ETA fusion |
|--|----|-------|------|------|---|--------|----------------------|
| | 45 | 348.2 | 48.6 | 876 | 1 | V10390 | Monoclonal antibody |

ALIGNMENTS

RESULT 1

786309 standard; DNA; 717 BP.

AC T86309;

DT 06-APR-1998 (first entry)

DE Single chain anti-disialoganglioside GD2 antibody 5f11-scFv.

KW Antibody construct; disialoganglioside; GD2; single chain Fv fragment

KW serv; tumour; neuroblastoma; osteosarcoma; soft tissue sarcoma;

KW tissue imaging; target delivery; toxin; streptavidin;

KW pro-drug converting enzyme; GD2-targeted lymphocyte; ss.

OS Synthetic.

PN W09734634-A1.

PD 25-SEP-1997.

PF 20-MAR-1997; U04427.

PR 20-MAR-1996; US 013703.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PI Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M;

DR WFL; 97-479996/44.

PT Recombinant single chain anti-disialoganglioside GD2 antibody -

PT useful to detect tumour cells expressing GD2 and to target

PT therapeutic agents, e.g. toxins, to such cells

PS Disclosure; page 11; 31pp; English.

CC The present sequence encodes a recombinant single chain peptide,

CC 5f11-scFv. The peptide is an antibody construct comprising the variab

CC regions of the heavy and light chains of an antibody against

CC disialoganglioside (GD2) as a single chain Fv fragment (scFv). GD2 oc

CC in many tumours types including neuroblastoma, osteosarcomas and o

CC soft tissue sarcomas, medulloblastomas, high grade astrocytomas,

CC melanomas and small cell lung cancer. The peptide can be detectably

CC labelled, preferably with 99m-Tc, for tissue imaging of cells express

CC GD2. It can also be used to target delivery of a therapeutic or

CC pre-therapeutic agent, such as a toxin, streptavidin or a pro-drug

CC converting enzyme, to cells expressing GD2. The peptide may further

CC comprise CD8 to facilitate the formation of GD2-targeted lymphocytes.

CC T cells containing the peptide can also be used to target GD2-produci

CC tumour cells.

CC Sequence 717 BP: 180 A: 181 C: 195 G: 159T:

| | Query Match | 99.7%; | Score 715; | DB 1; | Length 717; |
|----|-----------------------|--|---------------|-----------|-------------|
| | Best Local Similarity | 100.0%; | Pred. No. | 5.2e-196; | |
| | Matches 717; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps |
| Qy | 1 | caggtgaactgcagcagctcagacgacctgaactggtnagccctggggcttcagtgaagata | 60 | | |
| Db | 1 | CAGGTGAACCTGCAGCAGCTCAGGACCTGAACCTGGTGNAGCCTGGGGCTTCAGTGAAAGATA | 60 | | |
| Qy | 61 | tccgtcgaaagacttctgtgnacaatttcactgtaatacaccatgcactcgggtgaagcagagc | 120 | | |
| Db | 61 | TCCGTCAAGACTTCTGGANACAATTCACUGAATACACCATGCACATGGGTGAAGCAGAGC | 120 | | |
| Qy | 121 | catggaaagacctgagtgaggattgaggtgtattaatcctaacaatggttggtactaacac | 180 | | |
| Db | 121 | CATGGAAGAAGACCTTGAGTGGATTGGAGGTATTAAATCTAACAAATGGTGTACTAACTAC | 180 | | |
| Qy | 181 | aagcagaagttcaaggccaagggccaattgactgtagacaaagtcctcccagcacagccctac | 240 | | |
| Db | 181 | AAGCAGAAGTTCTAGGCCAAGGCCAATTGACTGTAGACAAGTCTCCAGCACAGCCTAC | 240 | | |
| Qy | 241 | atggagctcccgagcctgcacatctgaggtctcgagttctgtacttgtaccaagatatct | 300 | | |
| Db | 241 | ATGGAGCTCCCGAGCCCTGACATCTGAGGATCTGCAGTCTATTACTGTGCAGAGAGATCT | 300 | | |
| Qy | 301 | acgggtcccgcttgcttactgggtgccaaaggagaccaggtcacccgctccctcaggtggagnc | 360 | | |
| Db | 301 | ACGGTCCCGCTTGCTTACTTGSSTTCCAAAGGACCATCGGTACCGTCTCCCTCAGGTGGAGC | 360 | | |

DR WPI: 99-243623/20.
DR P-PSDB; Y05363.
PT Single-stranded antibody against hepatitis B virus core protein,
PT applicable as (gene) therapeutic agents for treatment of hepatitis B
PT viral infections
PS Claim 5; Page 55-57; 72pp; English.
CC This sequence encodes the single-stranded antibody of the invention, that
CC has the capability of binding to a hepatitis B virus (HBV) core protein.
CC Therapeutic agents can be formulated with the antibody for treatment of
CC HBV infections by stopping proliferation of the virus through inhibition
CC of viral DNA synthesis, and the gene encoding the antibody can be applied
CC as an agent to gene therapy.
SQ Sequence 394 BP; 217 A; 225 C; 237 G; 215 T;

Query Match 73.1%; Score 523.8; DB 1; Length 894;
Best Local Similarity 85.6%; Pred. No. 4e-141;
Matches 624; Conservative 0; Mismatches 84; Indels 21; Gaps 3;
QY 1 caggtgaactgcagcagctcagcagcctgaaactggtgagcctggggcttcagtgaaagata 60
Db 118 CAGGTGAAGCTGCAGGAGTGCAGGAGCTGAGCTGGAGAGGCTGGCGCTTCAGTGAAGATA 177
QY 61 tctctgaagactctctgaanacaaattcactgaatacaccactgactggtgaaagcagc 120
Db 178 TCCTGCAAGGCTTCGTGTTACTCAATTCAGTGGCTACAAATGAAGTGGGGAACACAGC 237
QY 121 catggaagagccttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtc 180
Db 238 AATGGAAGAGCCTTGAGTGGATGGATATATTTATCTTACATGTTGTTACTGGCTAC 297
QY 181 aagcagagtgtaagggcagcagccacattgactgtagacaaagtcctccagcagcctac 240
Db 298 AACCAAGAGTTCAGAGCAAGGCGCACATTGACTGTAGACAAATCTCCAGCACAGCCTAC 357
QY 241 atgagctcgcagcctgcacatctgagattctgagattctgagattctgagattctgagattct 300
Db 358 ATGCAACTGAGCAGCCTGACATCTGAGGACTCTGAGTCTGATCTGATCTGATCTGAGG 412
QY 301 acggtccctgttctactggttccagggagcagcagcagcagcagcagcagcagcagcagc 360
Db 412 --CTGGGACTTGACTACTGGGGCAAGGGAGCACAGGTCACCGTCTCTCAGGTGGAGGC 468
QY 361 ggttcagggagagtggtctctggtggtggtggtggtggtggtggtggtggtggtggttcca 420
Db 469 GGTTCAGGCGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCACTCTCCA 528
QY 421 gcaatcatgtctacatctccagggagagaggtccacatgacatgacatgacatgacatgacat 480
Db 529 ACCACATGCTGATCTCCGGGGGAGAGATCACTATCACCTGCACTGCACTGCACTCAAGT 588
QY 481 atag-----ttacatgactgtagcagcagcagcagcagcagcagcagcagcagcagcagc 534
Db 589 ATAAGTTCCAATTTACTTGTGATGATACAGAGAAGCCAGGATTCCTCCCTAAACTCTTG 648
QY 535 attatgacacatccaaactggtctctggtggtggtggtggtggtggtggtggtggtggttct 594
Db 649 ATTTATAGGACATCAATCTGGCTCTGGAAATCCAGCTCGCTTCAGTGGCAGTGGGTCT 708
QY 595 gggacaccttattctcacaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 654
Db 709 GGGACCTTCTACTCTCACAAATGGACCAATGGAGGCTGAAGATGTGGCCACTTACTAC 768
QY 655 tgccatcagcagagtagttaccgc-----tcacgttctggtggtggtggtggtggtggtggtg 708
Db 769 TGCCAGCAGGAGTGTAGTATACCGCATATTCACGTTCCGTTCCGTTCCGTTCCGTTCCGAA 828
QY 709 ataaacagg 717
Db 829 ATRAAACGG 837

RESULT 4

T04023
ID T04023 standard; cDNA; 726 BP.
AC T04023;
DT 02-JUL-1996 (first entry)
DE Anti-EGFR single chain antibody (Clone 5 F 1).
KW Single chain antibody; antibody; epidermal growth factor receptor;
KW EGFR; tumour; cancer; glioma; melanoma; carcinoma; diagnosis;
OS Mus musculus.
FH Key Location/Qualifiers
FT 1..726
FT cds
FT /tag- a
FT /product= Single-chain Fv, heavy and light chain
FT plus linker.
PN WO9525167-A1.
PD 21-SEP-1995.
PF 16-MAR-1995; E00978.
PR 17-MAR-1994; EP-104160.
PR 02-DEC-1994; EP-118970.
PA (MERE) MERCK PATENT GMBH.
PI Adan J, Ansell KH, Bendig MM, Blasco F, Guessow D;
PI Kettleborough AC, Mitjans F, Piulats J, Rosell E;
DR WPI: 95-336972/43.
DR P-PSDB: R79870.
PT Anti-EGFR antibodies and single chain Fv antibody fragments -
PT obtained from phage-antibody libraries, useful for diagnosis and
PT therapy of tumours
PS Claim 4; Page 66-67; 93pp; English.
CC Anti-epidermal growth factor receptor (EGFR) single chain antibodies
CC and antibodies constructed from anti-EGFR antibody fragments can be
CC used for diagnosis of tumours and assessment of tumour growth in
CC vitro and in vivo. They may also be used in a pharmaceutical
CC composition for the therapy of e.g. melanomas, gliomas and carcinomas.
CC The antibodies and fragments are derived from mice but are humanised
CC so as to cause minimum reaction against them. They are produced
CC using the phage antibody library. (See T04011-T04026 and
CC R79858-R79873)
SQ Sequence 726 BP; 174 A; 192 C; 206 G; 154 T;

Query Match 72.7%; Score 521.6; DB 1; Length 726;
Best Local Similarity 84.4%; Pred. No. 1.6e-140;
Matches 613; Conservative 0; Mismatches 101; Indels 12; Gaps 2;

QY 1 caggtgaactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 60
Db 1 CAGGTGAAGCTGCAGGAGCTGGGGCTGAACCTGGAAGCCTGGGCTTCAGTGAAGTTG 60
QY 61 tctctgaagactctggaanacaaattcactgaaatcacacacactgactggtggtgagcagc 120
Db 61 TCCTGCAAGGCTTCGGGCTACACCTTACAGCCACTGGATGCTGGGTGAAGCAGAGG 120
QY 121 catggaagagccttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtc 180
Db 121 GCTGGACAGGCTTGAGTGGATCGGAGAGATTAATCCAGAACGGCGCTTACTACTAC 180
QY 181 aagcagaagtctcaaggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
Db 181 AATGGAATTCAGAGCAAGGCGCACATGCTGTAGACAAATCTCCAGCACAGCCTAC 240
QY 241 atggaactcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 299
Db 241 ATGCAACTCAGCAGCTGACATCTGAGGACTCTGGGCTCTATTACTGTGCCAGTCGGGAC 300
QY 300 tacggtcccg-----tttgcctactggttccaaaggcagcagcagcagcagcagcagc 348
Db 301 TATGATTACAGCAGGAGCTGACTTTGACTGCGGCCAAGGGACACAGGTCACCTCTCC 360
QY 349 tcaggtggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 408
Db 361 TCAGGTGGCGGTGGCTCGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 420
QY 409 actcagttctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 468

CC used in the antibody-firefly luciferase fusion protein of the invention
CC in which an antibody part consisting of a peptide having antibody
CC activity is combined with an enzyme part consisting of firefly
CC luciferase.
SQ Sequence 732 BP; 179 A; 183 C; 212 G; 158 T

| | | | |
|-----------------------|--------------|-----------------|---------|
| Query Match | 72.6% | Score 520.4; | DB 1; |
| Best Local Similarity | 84.5%; | | |
| Matches 610; | Conservative | | |
| | 0; | Mismatches 103; | |
| | | Indels 9; | Gaps 2; |
| | | Length 732; | |

[illegible][illegible][illegible]

50 246 1GCAGGC1CAGCAGGCC1AACAT1C1GAGAGAC1CTGCGG1CAT111C1G1ACAAAGA...--A1GG 300
 Qy 302 CGTCCCGTTTGCTACTGGTCCAGGGGACCAAGTCACCGTCTCCTCCTCAGGTGGAGCG 361

| Accession | Sequence | Length |
|-----------|---|--------|
| Db | 305 GGSTACGACGGGACTACTGGGEGCCAAAGGACCACGGTCACCGTCTCCTCAGGTGGAGCGC | 364 |
| Qy | 362 gttcaggcgaggtggctctggcggtggcgagatcggacatcgagctcactcagctccag | 421 |

365 GTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCAGTCTCCAG 424

Db

422 caatcatgtctgcattctccagggggaagaaggtcaccaagacctgcagtggaagctcaagta 481

Oy

Db 425 CAATCATGTCTGCATCTCTAGGGGAACGGGTACCACCTGACCTGCCAGTCAAGTG 484

482 caagc...acacgacacggtaccagcagaagccgctacgtcccccacaaagatgga 533
 485 TRAGTCCGCAACTTGCACTGGTACCAGCAGAAGTCAGAAACGTCCCCAACCCCTTGA 544

| | | | |
|----|-----|---|-----|
| Qy | 536 | tttatgaacatccaaactggcttctggagtcocctgcctcaagtgcagtgggctg | 595 |
| | | | |
| Db | 545 | TTTATGGACATCCAACTGGCTTCTGGAATCCCTGTCGTTTCAGTGGCAGTGGATCTG | 604 |

Qy 596 ggaccttatttctctcaaatcagcagcatggaggctgtagatgctgcacattattact 655
|||||
Db 605 GGACCTTATTCTCTCAATCAGCAGCATGGAGGCTGAGATGCTGCACATTATTACT 664
|||||

QY 656 gccatcagcggagtagttaccgcgtcacgttcgggtcgtgggacacacagttgggaaataaaac 715

DB 665 GCCAGCAGTGGAGTAGTAAACCCATTACAGTTCGGGCTCGGGGACCAAGCTGGAAATAAAAC 721

QY 716 gg 717
||

Db 725 GG 726

RESULT 7
V23580
ID V23580 standard; cDNA to mRNA; 2364 BP.
AC V23580:

DE Anticapsid-Firefly luciferase fusion protein gene.
KW Firefly luciferase; antibody-luciferase fusion protein; ds.

| OS | Luciola cruciata. |
|----|---------------------|
| FH | Key |
| | Location/Qualifiers |

```

CDS
    1. .2364
    /*tag= a
    /transl_except= (pos: 670.. 672, aa: Glu)
    /transl_except= (pos: 739..741, aa: Trp)
    /transl_except= (pos: 1369..1371, aa: Ala)
    /note= "no stop codon given"

J09187281-A.
22-JUL-1997.
09-JAN-1996; 001812.
09-JAN-1996; JP-001812.
(KIKK ) KIKKOMAN CORP.
WPI: 98-275089/25.
P-PSDB: W53882.
DR P-PSDB: W53882.
DR Antibody-firfly luciferase fused protein - and related products
PT i.e. firfly luciferase fused gene, recombinant DNA and its
PT preparation
PS Disclosure: Page 13: 17pp; Japanese.
CC This sequence encodes the fusion protein of the invention. The protein
CC a antibody-firfly luciferase fusion protein, in which an antibody pa
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firfly luciferase.
CC Sequence 2364 BP: 697 A; 559 G; 664 T;
SQ

```

Query Match 72.2%; Score 517.8; DB 1; Length 2364;
Best Local Similarity 84.3%; Pred. No. 3e-139;
Matches 608; Conservative 0; Mismatches 104; Indels 9; Gaps

Qy 2 aggtgaactgcagcagctcaggacctgaactggtgnagcctggggcttcagtgagatat 61
|||||
dh 8 AGCTGAAGTTCAGCAGTCAAGGAGTCAGCTGCTGAGGCTTGGCTCCCTACCTGACATT 67

Qy 62 cctgcaagacttctgganacaaattcactgaatacaccatgcactgggtgaagcagagcc 121

Db 68 CCTGCAAGGCTTCTGGCTATGCTATTGCTAGTACTGATGAAGCTGGGTGAAGCAGAGGC 127

QY 122 atggaagaagagcccttgagtggaattggagggtattaatacctaacaatgggtggtactaactaca 181

Db 128 CTGGACAGGCTCTCAGTGGATTGGACAGATTAACTCTGGAGATGGTGATACGAATTACA 187

Qy 182 agcagaagttaagggcaaggccacattgactgtagacaagtcctccgacacagcctaca 241

Db 188 ATGGAAGTTCAAGGGTAAAGTCACACTGACTGCAGACAAATCCTCCAGCACAGCCCACA 247

Ov 242 tccagactcaccagacgtgacatctcgaggattctgcagtttatattactgtgccaaataactaca 301

Q3
248 TGCAGCTCAGCAGCCTAACATCTGAGACTCTGCGGTCTATTTCTGTACAAGA---ATGG 304
Db

Qy 302 cgggtcccgattgcttacttggtccaagggaacacggtaccggtcctcctcaggtgagggcg 361
||| | + ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 305 GGGTACGAGCGGGACTACTGGGGCCAAAGGGACCAGGTACCGTCTCCTCAAGTTGAGGC 364

QY 362 gttcaggcgagggtggctctcggcggtggcggtcgacatcgagctcactcagttccag 421
|||||
Db 365 GTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCAGTCTCCAG 424

Qy 422 caatcatgtctgcattctccagggggaaggtcaccatgcacctgcagctggcagctcaagta 481

QY . 482 taagtt-----acatgcactggtaccagcagaagcctgtcacctccccccaaagatgga 535
 |||||
 DB CAATCATGATCTGCAATCTCTAGGGGAAACGGGTACCATGACCTTGCACCTGCCAGCTCAAGTG 488

| | | | |
|----|-----|--|-----|
| Db | 485 | TAAGTTCACGAACATTGCATGGTACGACGAGAAGTCAGAAACCTCCCCAAACCTGGA | 544 |
| Qy | 536 | tttatgacacatccaaactggcttctggagtcacctgctcgcttcagtggcagtggggtctg | 595 |

Db 545 TTTATGGCACATCCAACTTGGCTTCTGGAAATCCCTGTTCCGCTTCAGTGGCAGTGGATCTG 604

Qv 596 ggacctcttattctcaccaatcagcagcatcgaggctctagatgctgcccatttact 655

605 GGACCTCTATTCTCACAATCAGCAGCATGGAGGGTGAAGATGCTGCCACTTATTACT 664

QY 656 gccatcagcggag-aagtTaccgcgtcacgttcggtgcctgggacacagtcggaaataaaac /T

| | | | | |
|----|---|--|--|--|
| PT | therapy of tumours | | | |
| PS | Claim 4; Page 68-69; 93pp: English. | | | |
| CC | Anti-epidermal growth factor receptor (EGFR) single chain antibodies | | | |
| CC | and antibodies constructed from anti-EGFR antibody fragments can be | | | |
| CC | used for diagnosis of tumours and assessment of tumour growth in | | | |
| CC | vitro and in vivo. They may also be used in a pharmaceutical | | | |
| CC | composition for the therapy of e.g. melanomas, gliomas and carcinomas | | | |
| CC | The antibodies and fragments are derived from mice but are humanised | | | |
| CC | so as to cause minimum reaction against them. They are produced | | | |
| CC | using the phage antibody library. (See T04011-T04026 and | | | |
| CC | R79858-R79873) | | | |
| CC | Sequence 726 BP; 171 A; 193 C; 205 G; 157 T; | | | |
| SQ | | | | |

Query Match 71.5%; Score 512.6; DB 1; Length 726;
Best Local Similarity 83.7%; Pred. No. 6.1e-138;
Matches 607; Conservative 0; Mismatches 106; Indels 12; Gaps

| | | | |
|----|-----|---|-----|
| QY | 2 | aggtgaaactcgacgagtcaggacctgaactcgtgtagcctgggggttcagtgaaagatat | 61 |
| Db | | | |
| Db | 2 | AGGTCAAGCTGCAGCAGTCAGGGGCTGAACGTGTGAAGCTGGGGCTTCAGTGAAGTTGT | 61 |
| QY | 62 | ctgtcaagactcttgaganacaaattcaactgaatatcacaccatgactcgggtgaagcagagcc | 121 |
| Db | | | |
| Db | 62 | CTGCAAGGCTTCGGGTACACCTTCACAGCCACTTGGATCACTGGGTGAAGCAGAGGG | 121 |
| QY | 122 | atggaaagagccttgagtggattggaggatttaattcctaacaattggttactaaactaca | 181 |
| Db | | | |
| Db | 122 | CTTGGCAAGGCCCTTGAGTGCATCGGACATTTAATCCACACAGCCGCTACTACTACTA | 181 |
| QY | 182 | agcagaagttcaaggcgaaggccacattgactgtagacaagtcctccacacagcctaca | 241 |
| Db | | | |
| Db | 182 | ATGAGAAATTCAGAGCAAGGCCACACTGACTGTAGACAAATCCTCCAGCAGCCCTACA | 241 |
| QY | 242 | tggagctcgcagccctgacatctgaagattctcgaagtctattactctgcaag-agatact | 300 |
| Db | | | |
| Db | 242 | TCGAACCTCAGCAGCCCTGACATCTGAGGACTGCTCGGTCTATTACTGTGCAGTCGGGACT | 301 |
| QY | 301 | acggtcccg-----tttgcttactgggtccaaaggacacaggtccacggtctcct | 349 |
| Db | | | |
| Db | 302 | ATGATTCAGCAGCGACGGTACTTTTGACTACTGGGCCCAAGGGACCCGGTCAACGCTCTCGT | 361 |
| QY | 350 | caggtggaggcgttccaggcggagggtggctctgctggcgtggcggaatcggacatcgagctca | 409 |
| Db | | | |
| Db | 362 | CAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGCGCGGATGTGACATTGAGCTCA | 421 |
| QY | 410 | ctcagttccagcaaatcatgtctgcattctccaggggagaaggtccacatgacctgcagtg | 469 |
| Db | | | |
| Db | 422 | CCCAGTCTCCAAACAATCATGCTGCAATCTCCAGGGAGAAGGTACCAATGACTTGCAGTG | 481 |
| QY | 470 | gcagctcaagataaagttacatgcactggtaccagcagaagcctgtccaactcccccaaa | 529 |
| Db | | | |
| Db | 482 | ACAGCTCAAGTGTAAATTACATGCTACTGTTACCCAGCAGAGCAGAGATCCTCCCCAGAC | 541 |
| QY | 530 | gatggatttatgacacatccaaactggctcttgagttccctgctcgtctcagtgggcagtg | 589 |
| Db | | | |
| Db | 542 | TTCTGATTTATGACATCCAACTGGCTTCTGGAGTCCGCTGCTCAGTGGGCAGTG | 601 |
| QY | 590 | ggtctgggaaccttattctctcaaatcagcagatggaggctgtagatgctgcacatt | 649 |
| Db | | | |
| Db | 602 | GGTCTGGGACCTTCTCTCTCACAAATCAGCCGAATGGAGGCTGAAGATGTGTCACATT | 661 |
| QY | 650 | attactgccatcagcgagtagttaccgctcagttctcggtgctgggacacagtttgaaa | 709 |
| Db | | | |
| Db | 662 | ATTACTGCCAGCAGTGGAGTAGTTACCCGCTCAGTTCGGTCTGGGACCAAGCTGGAAA | 721 |
| QY | 710 | taaaa | 714 |
| Db | | | |
| Db | 722 | TAATAA | 726 |

RESULT 10
Q90563

| | |
|----|---|
| ID | Q90663 standard; DNA; 810 BP. |
| AC | Q90663; |
| DT | 26-JAN-1996 (first entry) |
| DE | MFE-23 antibody coding sequence. |
| KW | Antibody; MFE-23; carcinoembryonic antigen; CEA; colorectal tumour; |
| KW | therapy; diabody; ds. |
| OS | Mus musculus. |
| FH | Key |
| FT | Location/Qualifiers |
| FT | FT misc_feature 79..438 |
| FT | FT misc_feature /*tag= a |
| FT | FT misc_feature /note= "variable heavy chain" |
| FT | FT misc_feature 439..483 |
| FT | FT misc_feature /*tag= b |
| FT | FT misc_feature /note= "linker region" |
| FT | FT misc_feature 484..801 |
| FT | FT misc_feature /*tag= c |
| FT | FT misc_feature /note= "variable light chain" |
| PN | W09515341-A1. |
| PD | 08-JUN-1995. |
| PF | 05-DEC-1994; G02658. |
| PR | 03-DEC-1993; GB-024807. |
| PA | (CANC-) CANCER RES CAMPAIGN TECHNOLOGY. |
| PI | Begent RHJ, Chester KA, Hawkins RE; |
| DR | WPI: 95-215234/28. |
| DR | P-PSDB; R75719. |
| PT | Antibody for carcinoembryonic antigen - for treatment and diagnosis |
| PT | of colorectal cancer |
| PS | Claim 3; Page 48-49; 72pp; English. |
| CC | This sequence represents the coding sequence for the MFE-23 antibody |
| CC | The encoded protein is an antibody against carcinoembryonic antigen |
| CC | (CEA). CEA is a marker antigen for cancer imaging and therapy. The |
| CC | MFE-23 antibody sequence was obtained using phage technology. In the |
| CC | process, mice were immunised with CEA. The antibody variable region |
| CC | genes obtained from these mice were then amplified from cDNA and cloned |
| CC | as a single chain Fv (scFv) into bacteriophage vectors, producing a |
| CC | library. The phages that bound to biotinylated CEA were selected and |
| CC | amplified, and this sequence (and the protein it encodes) were selected |
| CC | The MFE-23 antibody was found to have good specificity and affinity |
| CC | CEA, meaning that it can be used in targeted anti-tumour therapies. |
| CC | humanised antibody with the complementarity determining regions of MFE |
| CC | may be made by CDR grafting. The antibody may be used for the treatment |
| CC | by surgery or therapy of a colorectal tumour, or in the diagnosis of |
| CC | colorectal tumour. MFE-23 may also be used to make diabodies (bivalent |
| CC | or bispecific antibody fragments which bind to two different antigens |
| CC | and may be linked to an antitumour agent or a detectable label. |
| SQ | sequence 810 BP; 189 A; 220 C; 222 G; 179; |

| | | | | |
|-----------------------|--------------|---------------------|-----------------|-------------|
| Query Match | 70.7%; | Score 507.2; | DB 1; | Length 810; |
| Best Local Similarity | 82.2%; | Pred. NO. 2.3e-136; | | |
| Matches 597; | Conservative | 0; | Mismatches 120; | Indels 9; |
| Gaps 1; | | | | |

| | | | |
|----|-----|---|-----|
| Qy | 1 | cagtgaaactgcaagcagtcagagacccctgaaactggtgnagcctggggcttcaagtgaagata | 60 |
| | | | |
| Db | 79 | CAGTGAAACTGCACGAGCTCTGGGCAGAACTTGTGAGGTCAGGGACCTCAGTCAAGATTG | 138 |
| | | | |
| Qy | 61 | tcttgcagaactctctgganacaaattcaactgaatacaccaatgcacctgggtgaagcagagc | 120 |
| | | | |
| Db | 139 | TCCTGCACAGCTTCTGCCTTCAACATTAAAGACTCTCTATATGACATGTTGAGCGCAGGG | 198 |
| | | | |
| Qy | 121 | catgaaagagccttgagtgattgaggtattaaactctaaacaattggtgctaactaac | 180 |
| | | | |
| Db | 199 | CCTGAACAGCGCTCGAGTCGGATTGGATGGATTGATCCTGAGAATGGTGATCTGAATAT | 258 |
| | | | |
| Qy | 191 | aagcagaagttcaaggcgacagggcccaattgactgtgacagaagtccctccagcacagcctac | 240 |
| | | | |
| Db | 259 | GCCCCGAAGTTCAGGGCCAGGGCCACTTTTACTACAGACACATCTCCCAACACAGCCTAC | 318 |
| | | | |
| Qy | 241 | atgagactcgcagcctgcacatctgcagattcttcagcttattactgtgcagaagatact | 300 |
| | | | |
| Db | 319 | CTGCAGCTCAGACGCCCTGACATCTGAGGACCATCGCCGCTATTATTATGAGGGGACT | 378 |
| | | | |
| Qy | 301 | -----acggctcccgctttgcttactggtgtccaaagggaccacggtcacgcttctccta | 351 |

| | |
|--|---|
| PD | 21-SEP-1995. |
| PF | 16-MAR-1995; E00978. |
| PR | 17-MAR-1994; EP-104160. |
| PR | 02-DEC-1994; EP-118970. |
| PA | (MERE) MERCK PATENT GMBH. |
| PI | Adan J, Ansell KH, Bendig MM, Blasco F, Guessow D; |
| PI | Kettleborough AC, Mitjans F, Pluats J, Rosell E; |
| DR | WFI; 95-336972/43. |
| DR | P-PSDB; R79873. |
| DR | Anti-EGFR antibodies and single chain fv antibody fragments - |
| PT | obtained from phage-antibody libraries, useful for diagnosis and |
| PT | therapy of tumours |
| PS | Claim 4; Page 73-74; 93pp; English. |
| CC | Anti-epidermal growth factor receptor (EGFR) single chain antibodies |
| CC | and antibodies constructed from anti-EGFR antibody fragments can be |
| CC | used for diagnosis of tumours and assessment of tumour growth in |
| CC | vivo and in vivo. They may also be used in a pharmaceutical |
| CC | composition for the therapy of e.g. melanomas, gliomas and carcinomas. |
| CC | The antibodies and fragments are derived from mice but are humanised |
| CC | so as to cause minimum reaction against them. they are produced |
| CC | using the phage antibody library. (See T04011-T04026 and |
| CC | R79858-R79873) |
| SQ | Sequence 732 BP; 175 A; 191 C; 209 G; 157 T; |
| | |
| Query Match | 70.4%; Score 504.6; DB 1; Length 732; |
| Best Local Similarity | 83.7%; Pred. No. 1.2e-135; |
| Matches 612; Conservative 0; Mismatches 101; Indels 18; Gaps | |
| QY | 2 aggtgaaactgcagcagtgcaggacctgaactggtgnagcttgggcttcagtgagaatat 61 |
| Dd | 2 AGGTGCAGCTGCACGACTCGGGCTGAACCTGTTGTAAGCCTGGGGCTTCAGTCAAGTTGT 61 |
| QY | 62 cctgcaagactttcgganacaaattcactgaatcacaccatgcactgggtggaagcagagcc 121 |
| Dd | 62 CCTCAAGGGTTCCGGCTACACCTTCACCAGCCACTGGATGCATGGTGTAAGCAGAGGG 121 |
| QY | 122 atggaagaagccttgatggatggagggttaatacctaactaatcggttgtactaacata 181 |
| Dd | 122 CTGGACAAGGCCCTTGACTGGATCGGAGAGTTTTAATCCCAGCAACGCCCGTACTAACTACA 181 |
| QY | 182 acgagaagttcaaagggccaagccacattgacttagacaagtcctccagcacagcctaca 241 |
| Dd | 182 ATGGAATAATTCAGAGCAAGGCCACACTGCTAGTAGCAAAATCCTCCAGCACAGCTTACA 241 |
| QY | 242 tggagctccgcagcctgacatctgaggattctcgactctattactgtggaag-agatact 300 |

| | | | | |
|-----------------------|-----------------|---------------------|------------|-------------|
| Query Match | 70.4%; | Score 504.6; | DB 1; | Length 732; |
| Best Local Similarity | 83.7%; | Pred. No. 1.2e-135; | | |
| Matches 612; | Conservative 0; | Mismatches 101; | Indels 18; | Gaps |

| | | | |
|----|-----|--|-----|
| QY | 2 | aggtgaactgcagcagctcaggacctgaactgctgmagcctggggcttcagtgaaagatat | 61 |
| | | | |
| | | | |
| Db | 2 | AGGTGCACCTGTCAGCAGCTCTGGGGCTGAACCTGGTGAAGCCTGGGGCTTCAGTGAAGTTGT | 61 |
| | | | |
| | | | |
| QY | 62 | cctgcaagactcttgganacaaatttcaactgaatacacccatgcactgggtggaagcagagcc | 121 |
| | | | |
| | | | |
| Db | 62 | CCTGCAAGGCTTCGGGTACACCTTCACCGCCACTGGATGCACCTGGGTGAAGCAGAGGG | 121 |
| | | | |
| | | | |
| QY | 122 | atggaagaagccttgatggatggagggtattaatcctcaacaatggtgtaactaataca | 181 |
| | | | |
| | | | |
| Db | 122 | CTGACAAAGCCCTTGATGGATCGGAGAGTTTTAATCCGACACGCCGCTACTAACTACA | 181 |
| | | | |
| | | | |
| QY | 182 | agcagaagttcgaagggcaagggcacattgactgtagacaagtcctccagcacagcctaca | 241 |
| | | | |
| | | | |
| Db | 182 | ATGAGAAATTCAGAGCAAGCCACACTGACTGTAGACAAATCCTCCAGCACAGCTTACA | 241 |
| | | | |
| | | | |
| QY | 242 | tgaagctccagcctgcacatctgaggattctgcagttctattactgtgcaag-agatact | 300 |
| | | | |
| | | | |
| Db | 242 | TGCAACTTCAGCAGCCTGACATCTGAGGACTCTCGGTCTATTACTGTGCACGTCGGGACT | 301 |
| | | | |
| | | | |
| QY | 301 | acggtcccg-----ttgttactgggtccaaggacacaggttcacccgtctctct | 349 |
| | | | |
| | | | |
| Db | 302 | ATGATTCAGCGGACGGTACTTTGACTACTTGGGCCAAGGGACACGGTCACCGTCTCCT | 361 |
| | | | |
| | | | |
| QY | 350 | caggtggaggggttcagggcgaggtggctctggcggtgcgggatcgacatcgagctca | 409 |
| | | | |
| | | | |
| Db | 362 | CAGGTGCGGTGGCTCGGGCGGTGGTGGTGGGTGGCGCGGATCTGACATTGAGCTCA | 421 |
| | | | |
| | | | |
| QY | 410 | ctcagttccagcaaatatgtctgcattctccaggggagaaggtcaacatgacctgcagtg | 469 |
| | | | |
| | | | |
| Db | 422 | CCCAGTCTCCAAACAATCATGTCTGCATCTCCAGGGAGAAGGTCACCATGACCTGCAGTG | 481 |
| | | | |
| | | | |
| QY | 470 | gcagctcaagataaagttacatgcactggtaaccagcagaagcctgtcaacctccccaaaa | 529 |
| | | | |
| | | | |
| Db | 482 | ACAGCTCAAGTGAAGTTACATGTACTGGTACCAGAGAGACAGGATCCTCCCCCAGAC | 541 |
| | | | |
| | | | |
| QY | 530 | gatggattatgacacatccaaactggcttcggagtcctctgccttcagtggcagtg | 589 |
| | | | |
| | | | |
| Db | 542 | TCCTGATTTATGACACATCCAACTGGCTTCTGGAGTCCCTGTTCCGTTCAAGTGGCAGTG | 601 |
| | | | |
| | | | |
| QY | 590 | ggctcgggaaccttattctctcacaatcagcagcatggaggctgtagatgctgcacatt | 649 |
| | | | |
| | | | |
| Db | 602 | GGTCTGGGACCTTACTCTCTCACAAATCAGCCGAATGGAGGCTGAAGATGCTGCCACTT | 661 |
| | | | |
| | | | |

```

Qy 650 attactgccatcagcgagtagttatccc-----gtccaggtctcggtctgggacacagt 703
      ||||| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 ATTATCGCAGCAGTGGAGTAGTATTACCCACCCATGTACACGTTTCGGAGGGGGACAAAGT 721
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 704 tggaaataaaaa 714
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 TGGAAATAAAAA 732
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
T04021
ID T04021 standard; cDNA; 732 BP.
AC T04021
DE 02-JUL-1996 (first entry)
DT Anti-EGFR single chain antibody (Clone 3 D 3).
DS Single chain antibody; antibody; epidermal growth factor receptor;
KW EGFR; tumour; cancer; glioma; melanoma; carcinoma; diagnosis;
KW assessment; phage-antibody library; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds J. .732
FT FT /*tag= a
FT FT /product= Single-chain Fv, heavy and light chain
FT FT plus linker.

WO9525167-A1.
PN 21-SEP-1995.
PD 16-MAR-1995; E00978.
PF 17-MAR-1994; EP-104160.
PR 02-DEC-1994; EP-118970.
PA (MERE ) MERCK PATENT GMBH.
PI Adan J, Ansell KH, Bendig MM, Blasco F, Guesow D;
PI Kettlborough AC, Mitjans F, Pluats J, Rosell E;
PI WPI; 95-336972/43.
DR P-PSDB: R79868.
DR Anti-EGFR antibodies and single chain Fv antibody fragments -
PT obtained from phage-antibody libraries, useful for diagnosis and
PT therapy of tumours
PT Claim 4; Page 61-62; 93pp; English.
PS Anti-epidermal growth factor receptor (EGFR) single chain antibodies
CC and antibodies constructed from anti-EGFR antibody fragments can be
CC used for diagnosis of tumours and assessment of tumour growth in
CC vitro and in vivo. They may also be used in a pharmaceutical
CC composition for the therapy of e.g. melanomas, gliomas and carcinomas.
CC The antibodies and fragments are derived from mice but are humanised
CC so as to cause minimum reaction against them. They are produced
CC using the phage antibody library. (See T04011-T04026 and
CC R79858-R79873)
SQ sequence 732 BP; 177 A; 195 C; 207 G; 153T;

```

| | Query Match | 70.2%; | Score 503; | DB 1; | Length 732; |
|----|-----------------------|---|---------------------|------------|-------------|
| | Best Local Similarity | 83.6%; | Pred. No. 3.5e-135; | | |
| | Matches 611; | Conservative 0; | Mismatches 102; | Indels 18; | Gaps 3; |
| Qy | 2 | agggtgaaactcagcagctcaggacctgaactggctggnagccctggggcttcagtgaaagatat | 61 | | |
| Db | 2 | AGGTCCTCAACTCGACGAGTCAGGGGCTGAACCTGGTGAAGCCTGGGGCTTCAGTGAAGTTGT | 61 | | |
| Qy | 62 | ctctcgaagactctctgganacaaattcactgaatcacacatgcactcggttgaaagcagagcc | 121 | | |
| Db | 62 | CCTGCAAGGCTTCGGGCTACACCTTCACAGCCACTCGATGCTGGGTGAAGCAGAGGG | 121 | | |
| Qy | 122 | atgdaaagagccttgagtgagattgaggtatattacctaacaatgggtgactaaactaca | 181 | | |
| Db | 122 | CTGGACAAGGCCTTGAGTGGATCGGAGAGTTATCCACGCAACGCCGCTACTTAACATACA | 181 | | |
| Qy | 182 | agcagaagttcgaaggccaaagccacattgactgtagacaagtcctccagcacacgcctaca | 241 | | |
| Db | 182 | ATGAAAAATCAAGAGCAAGGCCACACTGCTGTGAGACAAATCTCCAGCACACGCTTACA | 241 | | |
| Qy | 242 | tggagctccgagcgcctgacatctcagagattctgagttctattactgtgcaag-agataact | 300 | | |
| Db | 242 | TGCAACTCAGCAGGCTGCACATCTGAGAGACTCTGGGCTCTATTACTGTGCCAGTCGGGACT | 301 | | |

PT contains antibody binding site and transmembrane domain of receptor,
useful for enhancing immune responses to disease
PS Example 1; Fig 12A; 69pp; English.
CC A nucleic acid sequence (T69218) is provided of hybridoma 2E12
CC sfv. It was obtd. as an HindIII/BglII fragment by PCR amplification
CC (see also T69233-24). It is used in retrovirus vectors such as
CC pLNC-2e1hlg3h1b7-1tm (see also T69216) and pLNC-2e1hlgcds8gpi
CC (see also T69217) that code for novel sfv molecules which have been
CC modified by connecting a transmembrane domain of a cell surface
CC receptor to the antigen binding site of the molecule. This creates
CC artificial ligands that stimulate adhesion between cells and
CC enhance co-stimulatory activity during an immune response against
CC disease. Such vectors can be used for gene therapy of e.g. cancer.
SQ Sequence 752 BP; 188 A; 182 C; 205 G; 177T; 177T.

| | | | | |
|-----------------------|-----------------|--|------------|-------------|
| Query Match | 58.0%; | Score 414.4; | DB 1; | Length 752; |
| Best Local Similarity | 76.9%; | Pred. No. 3.7e-105; | | |
| Matches 569; | Conservative 0; | Mismatches 141; | Indels 30; | Gaps |
| QY | 4 | atgtgatgacccagcactcccaaatctcgtctgtatcagcagagacacagggttaccata | 63 | |
| Db | 7 | ATTGTGCTCACCCAAATCTCCAGCTCTTTGGCTGTGTCTGTAGGTCAGAGAGCCACCATC | 66 | |
| QY | 64 | acctgcaaggccagtcagagctgtgagtaaatgatgtggct-----tggttaccaa | 111 | |
| Db | 67 | TCCTGCAGAGCCAGTAAAGTGTTAATATTATGTCACAAAGTTAATGCAGTGGTACCAA | 126 | |
| QY | 112 | cagaagcaggcgagctccgaagaactcgtgataactctgcatcccaatcgctacactgga | 171 | |
| Db | 127 | CAGAAACACGAGCAGCGACCCAAACCTCTATCTGCTGCAATCCAAACGTAGAAATCTGG | 186 | |
| QY | 172 | gtccctgatcgcttcactggcgagtcgatatggagcaggatttccacttcaccatcagcact | 231 | |
| Db | 187 | GTCCCTGCCAGGTTTAGTGGCAGTGGCTGTGGACAGACTTCAGCCTTCACATCCATCT | 246 | |
| QY | 232 | gtcagggtgtaagacctggcgagttttctgtcagc-----aggattatagctcg | 282 | |
| Db | 247 | GTGAGAGGAGTATATTCGAATGTATTTCTGTGACGAAAGTAGGAAGGTTCCTTTGGAG | 306 | |
| QY | 283 | ctcggaggggggacaaagctggaaataaa---aggctggagcggttcacggcggagtggc | 339 | |
| Db | 307 | TTGGGTGGAGGCCAAGCTGGGAATCAAAACGAGGTGGCGGTGGCTCGGCGGTGGTGGG | 366 | |
| QY | 340 | tctggcggtggtgcggatcgcaggtgcaggtgaaagagtcaggaacctggcctggcgcccc | 399 | |
| Db | 367 | TCGGGTGGCGCGGATCTCAGGTGCAGCTGAAGSAGTCAGGAGCTTGGCTGGTGGCGCC | 426 | |
| QY | 400 | tcacagagcctgtccatcaacttcacatgtctctgggttttcatacccaattatggtgta | 459 | |
| Db | 427 | TCACAGAGCCTGTCCATCATCATGACCGCTTCAGGGTTCTCATTAACCGGCTATGTTGA | 486 | |
| QY | 460 | cactgggttcgcagcctcccagaaagggtctggagtggtggcgagtaatatggcgctggt | 519 | |
| Db | 487 | AAC TGGGTTCGCCAGCCTCCAGAAAGGGTCTGGAGTGGCTGGGAATGATATGGGGTGA | 546 | |
| QY | 520 | ggaagcacaaattataattcggctcttatgtccagactgagcatcgcaaggacaaactcc | 579 | |
| Db | 547 | GGAGGCACAGACTATAATTTCAGCTCTCAAAATCCAGACTGAGCATCACCAAGGACAACTCC | 606 | |
| QY | 580 | aagagccaagtttcttaaaatgaacagctctgcaactgatgacacacccatgtaactac | 639 | |
| Db | 607 | AAGAGCCCAAGTTTCTTAAAAATGAACAGTCTGCAACATGATGACACAGCCAGATACTAC | 666 | |
| QY | 640 | tgtgcagctcgggggggtaactc-----ggctatgctttgactactcggggtcgaagga | 693 | |
| Db | 667 | TGTGCCAGAGATGGTTATAGTAACATTTCATATGTTATGGACTACTTGGGGTCAAGGA | 726 | |
| QY | 694 | acctcagtcacctgtctctc | 713 | |
| Db | 727 | ACCTCAGTCACCGTCTCCTC | 746 | |

| | | |
|--------|--|------------------------------|
| RESULT | 4 | |
| ID | Q51541 | standard; DNA; 729 BP. |
| AC | Q51541; | |
| DT | 23-MAY-1994 | (first entry) |
| DE | Coding sequence of polypeptide with affinity for fluorescein. | |
| KW | Monoclonal antibody; MAb; affinity; binding; antigen; diagnostics; | |
| KW | therapy; imaging; purification; biosensors; ss. | |
| OS | Synthetic. | |
| PN | US5260203-A. | |
| PD | 09-NOV-1993. | |
| PF | 02-SEP-1986; | 902971. |
| PR | 02-SEP-1986; | US-902971. |
| PR | 02-SEP-1987; | US-092110. |
| PR | 19-JAN-1989; | US-299617. |
| PR | 25-APR-1990; | US-512910. |
| PA | (ENZO-) ENZON INC. | |
| PI | Bird RE, Hardman K, Ladner RC; | |
| DR | WPI: 93-367875/46. | |
| DR | P-PSDB: R43880. | |
| PT | Single chain polypeptide for binding antigen - comprising light | |
| PT | and heavy chain antigen binding portions linked by peptide linker | |
| PS | Example 9; Figure 40; 78pp; English. | |
| CC | The single chain polypeptide is derived from the mature light and | |
| CC | heavy chains of a monoclonal antibody (MAb) and has affinity | |
| CC | for a given antigen (Fluorescein). It comprises a first | |
| CC | polypeptide comprising the antigen binding portion of of the light | |
| CC | chain variable region of an antibody and a second polypeptide | |
| CC | comprising the antigen binding portion of the heavy chain variable | |
| CC | region of an antibody and at least one peptide linker linking the | |
| CC | first and second polypeptide chains. The resulting single chain | |
| CC | polypeptide can be used in diagnostics, therapy | |
| CC | (in vivo and in vitro), imaging, purifications and biosensors. | |
| CC | This particular single chain binding molecule was designated | |
| CC | 18-2-3/TR59 and contains one linker peptide. | |
| QO | Sequence | 729 BP; 177 A; 186 C; 178 T; |

| Query Match | 57.2% | Score 408.4 | DB 1 | Length 729 |
|-----------------------|---|--------------------|-----------|------------|
| Best Local Similarity | 76.7% | Pred. No. 1.6e-103 | | |
| Matches 557 | Conservative 0 | Mismatches 151 | Indels 18 | Gaps 4 |
| Qy 4 | atttgatgaccagactcccaaatctcgtctgttatcagcagagacaggggtaccata | 63 | | |
| Db 7 | AATGTGTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGAAAGGTCCACATG | 66 | | |
| Qy 64 | acctgcaaggccagctagagtgtagt---aatgatgtgcttggtaccacaagaagcca | 120 | | |
| Db 67 | ACCTGTCAGGGCCACTCAAGTGTAAAGTTCACGTACTGTCACACTGGTACCACGACGAGTCA | 126 | | |
| Qy 121 | gggcagctctccgaactcgtgatatactctgcataccaatcgcatacctggagtcgccctgat | 180 | | |
| Db 127 | GGTGCTCCCCAAACTCTGGGTTTATGGCACATCCAACTTGGCTTCTGGAGTCCCTGCT | 186 | | |
| Qy 181 | cgttctactggcagtgatattggacaggattctcaactttcaacatacagcactgtgcagct | 240 | | |
| Db 187 | CGCTTCAGTGCGGTGGGCTCTGGGACCTCTTACTCTCTCACAATCAGCAGTGTGGAGGCT | 246 | | |
| Qy 241 | gaagacctggcagtttttctctgcacaggattatag-----ctgcctcggaggg | 291 | | |
| Db 247 | GAAGATGTCGCCACTTATTACTGCCAGCAGTACAGTGGTTACCCACATCAGCTCGGTGCT | 306 | | |
| Qy 292 | gggaccaagctggaaaataaaaggtggagggcggttcaggcggaggtggctctggcggtggc | 351 | | |
| Db 307 | GGGACCAAGCT---TAAAGAACTCGGTTCGTTTCTTCTGAACAGCTGGCTCAGTTTCGT | 363 | | |
| Qy 352 | ggatctcaggtgcaggtggaaggagtcaggaacctggcctggctggcgccctcacagacctg | 411 | | |
| Db 364 | TCCTCGGATGTCAGCTGAAGAGTTCAGGACCTGTCTCGTGGTGGCGCCCTCACAGGCTG | 423 | | |
| Qy 412 | tccatcactgtcactgtctctgggtttttcaataaccaattatggtgtacactgggttcgc | 471 | | |
| Db 424 | TCCATCACTTCGACTGCTCTGGGTTTTTCATTAAACCAACTATGGTGTACACTGGGTTCGC | 483 | | |


```

Db 544 TATAATTCAGCTCTCATGTCCAGATGAGCATCAACAAAGACAATTCACAGAGCCAAAGTT 603
Qy 592 tcttaaaataacagctctgaaactgatcacagccatgtactactgtgcca---gt 648
Db 604 TCTTAAAAATGAACAGCTCTGCAAAATGATGACAGAGCCATATACTACTGTGCGCAACGA 663
Qy 649 cgggggggtaactcagctctgtcttggactactcgggtcaaggaaacctcagtcaccgtc 708
Db 664 CTGGAACGAATCTTTTACTATGCTATGAGCTATTGGGTCAAGGAACCTCAGTCACCGTC 723
Qy 709 tectca 714
Db 724 TCCTAA 729

RESULT 8
ID T69216 standard; DNA; 1528 BP.
AC T69216:
DT 19-AUG-1997 (first entry)
DE Retrovirus vector pLNC-2el2hlglhB7-1tm encoding modified sfv.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; retrovirus; vector; pLNC-2el2hlglhB7-1tm; ss.
OS Synthetic.
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 1; Fig 10A-B; 59pp; English.
CC A nucleic acid sequence (T69216) is provided of a modified sfv
CC encoded by retrovirus vector pLNC-2el2hlglhB7-1tm. This comprises
CC a fusion between hybridoma 2E12 sfv (see also T69218), human IgG1
CC Fc portion (see also T69219) and the transmembrane domain from
CC human B7-1 (CD80) (see also T69220) in vector pLNC. This is an
CC example of novel sfv molecules that are modified by connecting a
CC transmembrane domain of a cell surface receptor to the antigen
CC binding site of the molecule. This creates artificial ligands that
CC stimulate adhesion between cells and enhance co-stimulatory
CC activity during an immune response against disease. The vectors
CC can be used for gene therapy of e.g. cancer.
SQ Sequence 1528 BP; 380 A; 434 C; 399 G; 315 T;

Query Match 56.3%; Score 401.8; DB 1; Length 1528;
Best Local Similarity 76.7%; Pred. No. 1.3e-101;
Matches 568; Conservative 0; Mismatches 142; Indels 31; Gaps 5;

Qy 4 attgtgatgaccagctcccaaatctctgtgtatcagcaggagacagggttaccata 63
Db 1 ATTGTGTCACCACTCACTCCAGCTCTTCTTGGCTGTCTCTAGGTGACAGACCCCACTC 60
Qy 64 acctcagggcagctcagctgtgagtaagtgtggtc-----tggtaccacaa 111
Db 61 TCCTGCAGAGCCAGTGAAGTGTGAATATTATGTCACAAGTTAATGAGTGGTACCA 120
Qy 112 cag-aagcagggcagctctccgaactcgtgatctactctgcataccatcgtcactgg 170
Db 121 CAGAAACCCAGAGCAGACCCACCAACTCCTCATCTCTGCTGCATCCACAGTAGATCTGG 180
Qy 171 agtccctcagctcctcagctgagctggatgggacggatttcacacatcagcagc 230
Db 181 GGTCTCCGCCAGGTTTGTAGTGGCAGTGGGTCTGGGACAGACTTCAGCCCTCAACATCC 240
Qy 231 tgtcagactaaagacccggcgactttattctgtcagc-----aggattatagctc 281
Db 241 TGTGAGGAGGATGATATTGCATGATTTCTGTCAGCAAAAGTAGGAAGGTTCTTTGGAC 300

```

```

Qy 282 gctcggaggggggaccagctcgggaaataaaa---ggtggagcggttcacgagcgagggtg 338
Db 301 GTTCGGTGGAGGACCAAGCTGGAAATCAAAAGGGGTGGCGGTGGCTCGGCGGTGGTGG 360
Qy 339 ctcctggcgggtggcgagctgcagggcgagggggaagagagtcagggacctggtcgtggtg 398
Db 361 GTCGGGTGGCGGGGATCTCTCAGGTGCAGTGCAGCTGAAGGAGTTCAGGACCTGGCCCT 420
Qy 399 ctacagagcctgtccatcactgtctcgtgggttttccatttaaccaattatggtgt 458
Db 421 CTCACAGAGCCTGTCCATCATCATACCGTCTCAGGGTCTCATTAACCGGCTATGTTGT 480
Qy 459 acactgggttcggcagcctccagaaagggtcgtggagtggctggagtaatatgggtgg 518
Db 481 AAACCTGGGTTCGCCAGCCTCCAGAAAGGGTCTGAGTGGGTGGGAATGATATGGGTGA 540
Qy 519 tgggaagcacaaaattataattcggctcttattgtccagactgagcagcaaggacaactc 578
Db 541 TGAAGCACAGACTATAATTTCAGCTCTCAAAATCCAGACTGAGCATCACCAAGGACAATC 600
Qy 579 caagagccaaagtttcttaaaatgaacagctctgcaaaactgatcacagcagcatgtacta 638
Db 601 CAAGAGCCAGTTTTCTTAAAAATGAACAGTCTGCARACTGATGACACAGCCAGCATACTA 660
Qy 639 ctgtgccagtcggggggttaactac-----ggctatgcttggactactgggttcaagg 692
Db 661 CTGTGCCAGAGATGTTATAGTAACCTTTTCATTACTATATGTTATGGACTACTGGGTCAAG 720
Qy 693 aacctcagtcacgtctcctc 713
Db 721 AACCTCAGTCACCGTCTCCTC 741

RESULT 9
T69217
ID T69217 standard; DNA; 1510 BP.
AC T69217:
DT 19-AUG-1997 (first entry)
DE Retrovirus vector pLNC-2el2hlglhG1CD58GPI encoding modified sfv.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; retrovirus; vector; pLNC-2el2hlglhG1CD58GPI; ss.
OS Synthetic.
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 2; Fig 11A-B; 69pp; English.
CC A nucleic acid sequence (T69217) is provided of a modified sfv
CC encoded by retrovirus vector pLNC-2el2hlglhG1CD58GPI. This comprises
CC a fusion between hybridoma 2E12 sfv (see also T69218), human IgG1
CC Fc portion (see also T69219) and the transmembrane domain from the
CC CD58 glyco-phosphoinositol (GPI) anchor (see also T69221) in vector
CC pLNC. This is an example of novel sfv molecules modified by
CC connecting a transmembrane domain of a cell surface receptor to the
CC antigen binding site of the molecule. This creates artificial
CC ligands that stimulate adhesion between cells and enhance co-
CC stimulatory activity during an immune response against disease.
CC Such vectors can be used for gene therapy of e.g. cancer.
SQ Sequence 1510 BP; 381 A; 422 C; 394 G; 313 T;

Query Match 56.3%; Score 401.8; DB 1; Length 1510;
Best Local Similarity 76.7%; Pred. No. 1.3e-101;
Matches 568; Conservative 0; Mismatches 142; Indels 31; Gaps 5;

Qy 4 attgtgatgaccagctcccaaatctctgtgtatcagcaggagacagggttaccata 63

```



```

Db 1 ATGTGCTACCCCAATCTCCAGTCTCTTTGGCTGTGTCTTAGTGCAGAGACCACCATC 60
QY 64 acctgcaaggccagtcagagtgtagtaaatgatgtggtc-----tggtaccaa 111
Db 61 TCCTGCAGAGCCAGTGAAGAGTGTTGAATATATGTACACAAGTTTAAATGCAAGTGTGTACCAA 120
QY 112 cag-aagccaggccagtcctccgaaactgctgatatatactctgcatccaatcgctacactgg 170
Db 121 CAGAAACCCAGGACAGACGCCCAACTCCCTCATCTCTGCTGCATCCCAACAGTAGAATCTGG 180
QY 171 agtccctgacgtctcactgagcagtgatggagcagatttcactttccatccacagcac 230
Db 181 GGTCCCTGCCAGGTTTGTAGTGGCAGTGGTCTGGACAGACTTCAGCTCAACATCCATCC 240
QY 231 tgtcagcgtgaagacccgagctgtattctctgcagc-----agattatagctc 281
Db 241 TGTGGAGGAGATGATATTGCAATGATTTCTGTACGCAAGTAGGAAGGTTCTTTGGAC 300
QY 282 gctcggaggggggacaaagctgaaataaaa---ggtggaggcggttcagcgaggtgg 338
Db 301 GTTCGGTGGAGGACCAAGCTGGAATCAACCGGGTGGCGGTGGCTCGGGCGGTGGTG 360
QY 339 ctctggcgtggcgagtcgcagctgcaggtgaaggagtcaggacctggcctgtggcgcc 398
Db 361 GTCGGGTGGCGGGGATCTCAGGTGCAGCTGAAGGAGTCAGGACCTGGCTGTGGCGCC 420
QY 399 ctcacagagcctgtccatcactgcactgtctctggtttttcattaccaaattatggtg 458
Db 421 CTCACAGAGCCTGTCCATCATGTCACCGTCTCAGGGTCTCATTAACCGGCTATGGTG 480
QY 459 acactgggttcgcagcctccaggaagggtctggagtggtcggtggagtaatatgggtgg 518
Db 481 AAACCTGGGTTCCGACGCTCCAGGAAGGCTCTGGAGTGGCTGGGAATGATATGGGGTGA 540
QY 519 tgaagcacaaaataataatcgcgtcttattgtccagactgagcagatcagcagacaaactc 578
Db 541 TGAAGCACAGACTATAAATTCAGCTCTCAATCCAGACTCAGACTCACCAAGGACAACTC 600
QY 579 caagagccaaagtcttctaaataaagacagctctgcaactgatgacacagccatgtacta 638
Db 601 CAAGGCCAAGTTTCTTAAATGAACAGCTCTGCAACTGTATGACACAGCCAGATACTA 660
QY 639 ctgtgcagtcg99ggggttaactac-----ggctatgctttggactactg99ggtcaag 692
Db 661 CTGTGCCAGAGATGGTTATAGTAACCTTTCATTACTATGTATGACTACTGTGGGTCAAGG 720
QY 693 aacctcaagtcacgtctctc 713
Db 721 AACCTCAGTCACCGCTCTCCTC 741

```

RESULT 10

```

T69222
ID T69222 standard; DNA; 824 BP.
AC T69222;
DE Murine anti-human CD28 hybridoma 9.3 sfv VL-VH gene fusion.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; hybridoma 9.3; CD28; ss.
OS Mus sp.
FH Key
FT cds Location/Qualifiers
FT 7..819
FT /tag= a
FT signal_peptide 7..66
FT /tag= b
FT /note= "native 9.3 signal"
FT 68..816
FT /tag= c
FT mat_peptide
FT /product= VL-VH fusion
FT 364..369
FT /tag= d
FT /note= "Jk1"
FT

```

```

FT mrna 781..819
FT /*tag= e
FT /note= "JH4"
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
DR P-PSDB: W16688.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 3; Fig 13; 69pp; English.
CC A nucleic acid sequence (T69212) codes for 9.3 sfv (W16688)
CC comprising a fusion between the light chain and heavy chain
CC variable regions of murine anti-human CD28 hybridoma 9.3. The
CC light chain V region is rearranged with a J gene homologous to
CC murine Jk2, and the heavy chain V gene is rearranged with a J gene
CC segment homologous to murine JH4. The construct was obt'd. by PCR
CC amplification (see also T69229-34) of the heavy and light chain
CC regions. Novel sfv molecules are modified by connecting a
CC transmembrane domain of a cell surface receptor to the antigen
CC binding site of the molecule. This creates artificial ligands that
CC stimulate adhesion between cells and enhance co-stimulatory
CC activity during an immune response against disease. Vectors
CC encoding the scFv can be used for the gene therapy of cancer.
SQ Sequence 824 BP; 191 A; 197 C; 232 G; 204 T;

```

Query Match 56.1%; Score 400.2; DB 1; Length 824;

Best Local Similarity 76.4%; Pred. No. 3.1e-101;

Matches 572; Conservative 0; Mismatches 136; Indels 39; Gaps 5;

```

QY 4 attgtgatgaccagactcccaaatctctgtatcagcaggagacaggtttaccata 63
Db 70 ATTGTGCTCACCACATCTCCAGCTTTCTTTGGCTGTCTCTTAGGGCAGAGCCACCATC 129
QY 64 acctgcaaggccagtcagagtgtagtaaatgatgtggtc-----tggtaccaa 111
Db 130 TCCTGCAGAGCCAGTGAAGAGTGTGNAATATTATGTCACAAGTTTAAATGCAAGTGTGTACCA 189
QY 112 cagaagccagggtcagctctccgaaactgctgtatatactctgcatccaatcgctacactgga 171
Db 190 CAGAAGCCAGGACAGCAGCCCAACTCCCTCACTTTGCTGCATCCCAACGCTAGAACTCTGG 249
QY 172 gtcctgtatcgtctcactggcagtgagatggagcaggtttcactttcaccatcagcact 231
Db 250 GTCCCTGCCAGGTTTAGTGGCAGTGGGTCTGGGACAAACTTCAGCCTCAACATCCATCCT 309
QY 232 gtgcaggtcgaagacactggcagtttattctgtcagc-----agattatagctcg 282
Db 310 GTGGACGAGATGATGTTGCAATGTATTCTGTACGAAAGTAGGAAGTTCCTTACAGC 369
QY 283 ctgcgaggggggaccagcagctggaataaaa-----ggtggaggcggtttcagcgga 333
Db 370 TTCGAGAGGGGGACCAAGCTGGAATAAAGCGGCTTCGGGTGGTGGCGTTCGAGGT 429
QY 334 ggtggtctcggcgggtggcgatcg-----caggtgcaggtgaagagagtcaggaacctggc 387
Db 430 GCGGTTTCAGCGGGCGGTGATCCCTGGCTCAGGTGCAGCTGAAGGAGTCAGAGACTGGC 489
QY 388 ctgggtggccctcacagagcctgccactcactgtctctgggttttcattaacc 447
Db 490 CTGGTGACCCCTCACAGAGCCTGCCATCACTTGTACTGTCTCTGGGTTTTCATTAAGC 549
QY 448 aattatggtgtacactgggttcgccagcctccaggaagggtctgtgagtggtggagta 507
Db 550 GACTATGGTGTTCATTTGGGTTCGCCAGTCTCCAGACAGGAGTGGAGTGGCGAGTA 609
QY 508 atatgggtcgtggaagcacaaattataattcgggtcttattgtccagactgagcatcagc 567

```

Db 610 ATATGGCTGGTGGAGGACGCAATTAATTCGGCTCTCAATGTCAGAAAGACATCACC 669
 Qy 568 aaggacaactccaagcgaagcttttcttaaaatgaacagctctgcaaaactgatgacaca 627
 Db 670 AAAGACAACCTCAAGGCGCAAGTTTCTTAAATAAGAGAGTCTCAAGCTGATGACACA 729
 Qy 628 gccatgtactactgtgcag---tcgggggggtaactacgcgtatgcgttttggaactctgg 684
 Db 730 GCGGTGTTACTGTGCCAGAGATAAGGGATACCTCTATTACTATCTATGGACTACTGG 789
 Qy 685 ggtcaagaacactcagtcacgcgtctccctc 713
 Db 790 GGTCAAGGAACCTCAGTCACCGCTCCCTC 818
 RESULT 11
 ID T13739 standard; DNA; 720 BP.
 AC T13739;
 DT 11-OCT-1996 (first entry)
 DE Single chain binding molecule 18-2-3/TRY202' DNA.
 KW Antibody engineering; single polypeptide chain binding molecule;
 KW single chain antibody; SCA; heavy chain; light chain;
 KW monoclonal antibody; MAb; immunoaffinity purification;
 KW 18-2-3/TRY202'; ss.
 OS Chimeric Mus sp;
 OS Chimeric synthetic.
 FH Key Location/Qualifiers
 FT cds 619..621
 FT /*tag= a
 FT /transl_except= (619..621, aa:Ile)
 FT 658..660
 FT /*tag= b
 FT /transl_except= (658..660, aa:Gly)
 PN US5534621-A.
 PD 09-JUL-1996.
 PF 02-SEP-1986; US-902971.
 PR 02-SEP-1986; US-902971.
 PR 02-SEP-1987; US-092110.
 PR 19-JAN-1989; US-299617.
 PR 25-APR-1990; US-512910.
 PR 01-APR-1993; US-040440.
 PR 06-JUN-1995; US-468992.
 PA (ENZO-) ENZON LABS INC.
 PI Bird RE, Hardman K, Ladner RC;
 DR WPI: 96-333309/33.
 DR P-PSDB: R99649.
 PT Immuno-purificn. using single binding chain molecule including
 PT antigen-binding parts of antibody light and heavy chain variable
 PT regions connected by a linker - is smaller, stabler and less
 PT expensive than complete antibodies
 PS Example 9; Fig 39; 78pp; English.
 CC A DNA construct (T13739) codes for single chain binding molecule
 CC 18-2-3/TRY202' (R99649), in which VL and VH regions of anti-
 CC fluorescein monoclonal antibody 18-2-3 are joined by a peptide,
 CC linker composed primarily of alternating Gly and Ser residues, with
 CC Glu and Lys residues to enhance solubility. The construct was
 CC expressed in E.coli. Soluble, folded chain antibody (SCA) was
 CC obtd. which was capable of exhibiting a biological binding
 CC activity equivalent in specificity and affinity to that of a
 CC monoclonal antibody. The SCA has the advantages of smaller size,
 CC greater stability and reduced cost.
 SQ Sequence 720 BP; 180 A; 184 C; 175 G; 181 T;

Query Match 55.8%; Score 398.4; DB 1; Length 720;
 Best Local Similarity 76.2%; Pred. No. 9.3e-101;
 Matches 553; Conservative 0; Mismatches 146; Indels 27; Gaps 4;

Qy 4 attgtgatgccacagctcccaaaattcctgtgtatcagcagagcaggggttacata 63
 Db 7 AATGTGCTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAGGTCAACATG 66

Qy 64 acctgcaaggccagtcagagtgtagt---aatgatgtggttggtaccacaagaagcca 120
 Db 67 ACCTGCAGGGCCAGCTCAAGTGTAAAGTTCAGTTACTTGCACTGGTACCAGCAGAAGTCA 126
 Qy 121 gggcagctcgcgaactgctgatactctgcacccaatcgctacacagtgagtcctgat 180
 Db 127 GGTGCTCCCCCAAACTCTGGGTTTATGGCACATCCAACTTGGCTTCTGGAGTCCCCTGT 186
 Qy 181 cgcttcactggcagtgatattggacggatttccacttccaccatcagcactgtgcaggt 240
 Db 187 CGCTTCACTGGCGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGTGTGAGGCT 245
 Qy 241 gaagacctggcagtttattctgtcagcaggattatag-----ctcgctggaggg 291
 Db 247 GAAGATGCTGCCACTTATTACTGCCAGCAGTACAGTGGTTACCCACTCAGCTTCGGTGT 306
 Qy 292 gggaccaagctggaataaaaaagggtggaggggttcagggcggagggtgctctggcgggtggc 351
 Db 307 GGGACCAAGCTTGAGCTGGAAGGTAAA-----TCTTCTGGTTCTGGTTCCGAA 354
 Qy 352 ggaatcgaggtgcaggtgaaggagtcagggactcaggtgctggtggcgcctcacagagctg 411
 Db 355 TCTAAATCTACTCAGCTGAGGAGTCAAGGACCTGCTGCTGGTGGGCCCTCTCACAGAGCTG 414
 Qy 412 tccatcactgtcactgtctctctgggttttccattaaaccaattatggtgtacactgggttcgc 471
 Db 415 TCCATCACTTGCACCTGCTCTGGGTTTTCATTAACCAACTATGGTGTACACTGGGTTTCG 474
 Qy 472 cagcctccaggaagggtcctggagctggagctggtggcgcctcacagagcacaagt 531
 Db 475 CAGCTCCAGGAAAGGCTCTGGAGTGGCTGGGAGTAATATGGGCTGGTGGAAACACAAAT 534
 Qy 532 tataatcgcgtcttattgtccagactcagcatcagcagcagcactccaagagcagaagt 591
 Db 535 TATAATTCAGCTCTCATGTCCAGACTCAGCATCAGCAAGAACAAATTCAGAGAGCCAAGTT 594
 Qy 592 ttcttaaaatgaacagctctgcaaaactgatgacacagccatgtactactgtgcca---gt 648
 Db 595 TTCTTAAAAATGAACAGTCTGCAAAATGATGACACAGCCATATATACTACTGTGCCAAACGA 654
 Qy 649 cgggggggtaactacggtctatgctttggactactctgggggtcaaggaaacctcagtcacgc 708
 Db 655 CTGAACGAATCTTTTACTATGTATGGACTATTGGGCTCAAGGAACCTCAGTCACCCGTC 714
 Qy 709 tcctca 714
 Db 715 TCCTAA 720
 RESULT 12
 T36463
 ID T36463 standard; DNA; 720 BP.
 AC T36463;
 DT 13-NOV-1996 (first entry)
 DE 18-2-3/TRY202' single chain binding protein-encoding DNA.
 KW Antibody engineering; monoclonal antibody; MAb;
 KW single chain antibody; immunocassay; fluorescein; 18-2-3/TRY202';
 KW single chain binding protein; ss.
 OS Chimeric Mus musculus;
 OS Chimeric synthetic.
 FH Key Location/Qualifiers
 FT cds 658..660
 FT /*tag= a
 FT /transl_except= (658..660, aa:Gly)
 PN US5518889-A.
 PD 21-MAY-1996.
 PF 02-SEP-1986; US-902971.
 PR 02-SEP-1986; US-902971.
 PR 02-SEP-1987; US-092110.
 PR 19-JAN-1989; US-299617.
 PR 25-APR-1990; US-512910.
 PR 01-APR-1993; US-040440.
 PR 06-JUN-1995; US-468998.

| | | | |
|-----|----|--|-----|
| 181 | QY | cgcttcactcggcagtgatgataatggagcggatttcacatttcaccatcagcactgtgcaggct | 240 |
| 187 | Db | CGCTTCAGTGGCGGTGGGCTCTGGAGCTCTTACTCTCACAATCAGCAGTGTGGAGGCT | 246 |
| 241 | QY | gaagacctggcagttatttctgtcagcaggattatag-----ctcgctcggagg | 291 |
| 247 | Db | GAAGATGCTGCCACTTATTACTGCCAGCAGTACAGTGGTTACCCACTCAGTTCGTGTGT | 306 |
| 292 | QY | gggaccaagctggaataaaaaaggtggagcgggttcaggcggaggtggctctcgtcgtggc | 351 |
| 307 | Db | GGGACCACAGCTTGAGCTGGAGGTAAA-----TCCTCTGGTTCTGGTTCCGAA | 354 |
| 352 | QY | ggatcgaggtgcgagtggaaggagtcaggacctggcgtggcggccctcacagagcctg | 411 |
| 355 | Db | TCATAATCTACTCAGCTGAAGGAGTCAGGACCTGTCTTGGTGGCGCCCTCAGAGCCTG | 414 |
| 412 | QY | tccatcactgcactgctctcgtgggttttcattaaaccaattatggtgtacactgggttcgc | 471 |
| 415 | Db | TCCATCAGTGCAGTCTCTCTGGGTTTTTCATTAACCAACTATGGTGTACACTGGGTTCGC | 474 |
| 472 | QY | cagctccaggaaaaggctctgtagtggctgggagtaatatggcgtggtgggaagcacaaat | 531 |
| 475 | Db | CAGCTCCAGGAAAGGGTCTGGAGTGGCTGGGAGTAATATGGCTGGTGGGAACACAAAT | 534 |
| 532 | QY | tataatcggtctcttatgtccagactgagcatcagcaaggacaactcccaagagccaagt | 591 |
| 535 | Db | TATAATTCAGTCTCATGTCCAGACTGAGCATCAGCAAGACAATTCCAAGAGCCCAAGTT | 594 |
| 592 | QY | ttcttaaaatgaacagctgtgcaaacctgatgacacagccatgtactactgtgcca---gt | 648 |
| 595 | Db | TTCCTTAAATGAACAGTCTGCAAAATGATGACACAGCCATATACTACTGTGCCAAACGA | 654 |
| 649 | QY | cggggggggttaactacggcctatgctttggactactggtgggtcaaggaaacctcagtcaccgct | 708 |
| 655 | Db | CTGGAACGAATCTTTTACATGCTATGGACTATTGGGGTCAAGGAACCTCAGTCAACGTC | 714 |

| | | | | | | |
|----|---|-----------------|---------------------|------------|-------------|--|
| | Query Match | 55.8%; | Score 398.2; | DB 1; | Length 737; | |
| | Best Local Similarity | 75.2%; | Pred. No. 1.1e-100; | | | |
| | Matches 544; | Conservative 0; | Mismatches 158; | Indels 21; | Gaps 3; | |
| QY | 4 attgtgatgacccagaactccaaattctcgttgtatcagcaggagacagggttaccata | 63 | | | | |
| | | | | | | |
| Db | 14 ATCGAGCTACCCAGTCTCCAGCCTCCCTTTCTCGTGCTGGGAGAAACTGTCAACCATC | 73 | | | | |
| QY | 64 acctgcaggcgagtcagagtgtgagttaatgatgtggcttggtaccacaagaagccaggg | 123 | | | | |
| | | | | | | |
| Db | 74 ACATGTCGAGCAAGTGGGAATTATTCACAATATTATTAGCATGGTATCAGCAGAAAACAGGA | 133 | | | | |

```

QY 124 cagctcccaaacgtgtgatatactctgcatccaatcgctacactggagtcctctgtcgc 183
DB 134 AAATCTCCTCAGCTCTGTGCTATTATACAACAACACCTTAGCAGATGGTGTGCCATCAAGG 193
QY 184 ttcaactggcagtgatattggcgggattttcactttcaccatcagcactgtcaggtcgaa 243
DB 194 TTCAGTGGCAGTGGATCAGACACAAATATTCTCAAGATCAACAGCTGCAACCTGAA 253
QY 244 gacctggcagtttattctgtcagcaggatt-----atagctcgctcggaggggg 294
DB 254 GATTTGGAGTTATTACTGTCAACATTTTGGAGTACTCCTCGGACGTTCCGTGGAGGG 313
QY 295 accaagctggaaataaaa---ggtggagcgggttcaggcggggtgtcctgtgctggtgc 351
DB 314 ACCAAGCTCGAGATCAAAAGGGGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGC 373
QY 352 ggaatccaggtcagtggaagagtcaggacctggcctggctggcgccctcacagagcctg 411
DB 374 GGATCGCAGGTGCAGCTGCAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTG 433
QY 412 tccatcacttgcaactgtctctgggttttcaataaccaattatggtgtacactgggttcgc 471
DB 434 TCATCACATGCACCTCTCAGGGTTCTCATTAACCGGCTATGGTGAACCTGGGTTCGC 493
QY 472 cagctcccgaaaggggtcgtggagtggtggagtaatatggctggtggaagcacaaat 531
DB 494 CAGCCTCCAGGAAAGGCTGGAGTGGCTGGGAATGATTGGGGTGTATGGAAACACAGAC 553
QY 532 tataattcggctcttattgtccagactgagcatcagcaggacactccaagagcgaagt 591
DB 554 TATAATTCAGCTCTCAAAATCCAGATGAGCATCACAGGACAACTCCAAGAGCCAAAGTT 613
QY 592 ttcttaaaatgaacagctctgcaaaactgatcacacagcagcatgtactactgtccagtcgg 651
DB 614 TTCTTAAATGAACAGTCTGCACATGATGACACAGCCAGGTACTGTGTCAGAGAG 673
QY 652 gggggttaactacggctatctttggactactcgtgggttaaggaaacctcagtcacgcctcc 711
DB 674 AGAGAT-----TATAGCTTGACTACTGGGGCCAAAGGACCAGCGTCAACCGTCTCC 724
QY 712 tca 714
DB 725 TCA 727

RESULT 14
T94606
ID T94606 standard; cDNA; 737 BP.
AC T94606;
DE EcorI-HindIII insert of pUR.4124 nucleotide sequence.
KW Multivalent antigen binding protein; variable domain; binding site;
OS Synthetic.
FH Key
FT Location/Qualifiers
FT 11..730
FT /*tag= a
FT /product= "VLLys-linker-VHlys"
FT mat_peptide
FT 11..334
FT /*tag= b
FT /note= "VLLys"
FT 335..379
FT /*tag= c
FT /note= "linker"
FT mat_peptide
FT 380..727
FT /*tag= d
FT /note= "VHlys"
FT W09738102-A1.
PD 16-OCT-1997.
PF 26-MAR-1997; E01609.
PR 04-APR-1996; EP-302412.
PA (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.

```

```

PI Davis PJ, Van DER LOGT CPE, Verhoeven ME;
DR WPI: 97-512719/47.
PT P-PSDB; W35561.
PT Multivalent antigen binding protein - comprises antibody variable
PT domains, linked by their association, where each associated variable
PT domain pair forms an antigen binding site
PS Example 3; Pages 42-43; 100pp; English.
CC This is the nucleotide sequence of EcorI-HindIII insert of pUR4124. It
CC contains the DNA encoding VHLys-linker-VLLys sequences. This is used in
CC the construction of a novel multivalent antigen binding protein. This
CC novel protein comprises 2 polypeptides, each comprising, in series, 3 or
CC more variable domains of an antibody heavy or light chain respectively,
CC which are linked by association of the respective heavy and light chain
CC variable domains, where each associated variable domain pair forms an
CC antigen binding site. The multivalent antigen binding protein can be used
CC in medicine, diagnosis, therapy, an immunoassay method or for
CC purification. It may be used in targeting a tumour cell with natural
CC killer cells and cytotoxic agents. It may be used to target cell killing
CC enzymes, e.g. oxidases and peroxidases, to a species with which is an
CC antigenic component of a dental plaque, such as S. sanguis or S. mutans.
CC The antigen binding protein may conveniently be purified straight from
CC the supernatant using conventional purification techniques. As the
CC protein is self-assembling, there is no need to purify individual
CC subunits prior to coupling as in existing techniques.
SQ Sequence 737 BP; 191 A; 179 C; 200 G; 167 T;

```

```

Query Match 55.8%; Score 398.2; DB 1: Length 737;
Best Local Similarity 75.2%; Pred. No. 1.1e-100;
Matches 544; Conservative 0; Mismatches 158; Indels 21; Gaps 3;
QY 4 attgtgatccacagactcccaaaattcctgtgtatcagcaggagacaggggttaccata 63
DB 14 ATCGAGCTCACCCAGTCTCCAGCCCTCCCTTTCTGCGTCTGTGGGAGAACTGTCAACATC 73
QY 64 acctcgaagcgcagtcagagtgtagtaagtgtggtgtgtgtgtgtgtgtgtgtgtgtgt 123
DB 74 ACATCTCGAGCAAGTGGGAATATTCCAAATTTATTAGCATGGTATCAGCAGAAACAGGGA 133
QY 124 cagctccgaaactcgtgatatactctgcacccaatcgctacactggagtcctcgtatcgc 183
DB 134 AAATCTCCTCAGCTCCTGGTCTATTATACACAACTTAGCAGATGGTGTGCCATCAAGG 193
QY 184 ttcaactggcagtgatattggcggagatttcaactttcaccatcagcactgtgcaggtcgaa 243
DB 194 TTCAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAGATCAACAGCTGCAACCTGAA 253
QY 244 gacctggcagtttattctgtcagcaggatt-----atagctcgtcggaggggg 294
DB 254 GATTTGGAGTTATTACTGTCAACATTTTGGAGTACTCCTCGGACGTTCCGTGGAGGG 313
QY 295 accaagctggaaataaaa---ggtggagcgggttcagcggaggtggtccttggcgtgcgc 351
DB 314 ACCAAGCTCGAGATCAACAGCGGGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGC 373
QY 352 ggaatccaggtcagtggaagagtcaggacctggcctgggtggcgccctcacagagcctg 411
DB 374 GGATCGCAGGTGCAGCTGCAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTG 433
QY 412 tccatcacttgcaactgtctctgggttttcaataaccaattatggtgtacactgggttcgc 471
DB 434 TCATCACATGCACCTCTCAGGGTTCTCATTAACCGGCTATGGTGAACCTGGGTTCGC 493
QY 472 cagctcccgaaaggggtcgtggagtggtggagtaatatggctggtggaagcacaaat 531
DB 494 CAGCCTCCAGGAAAGGCTGGAGTGGCTGGGAATGATTGGGGTGTATGGAAACACAGAC 553
QY 532 tataattcggctcttattgtccagactgagcatcagcaggacactccaagagcgaagt 591
DB 554 TATAATTCAGCTCTCAAAATCCAGATGAGCATCACAGGACAACTCCAAGAGCCAAAGTT 613
QY 592 ttcttaaaatgaacagctctgcaaaactgatcacacagcagcatgtactactgtccagtcgg 651
DB 614 TTCTTAAATGAACAGTCTGCACATGATGACACAGCCAGGTACTGTGTCAGAGAG 673
QY 652 gggggttaactacggctatctttggactactcgtgggttaaggaaacctcagtcacgcctcc 711
DB 674 AGAGAT-----TATAGCTTGACTACTGGGGCCAAAGGACCAGCGTCAACCGTCTCC 724
QY 712 tca 714
DB 725 TCA 727

```

```

Db 614 TTCTTAAATGACAGCTGTCACACTGATGACACAGCCAGGTACTACTGTGCCAGAG 673
QY 652 ggggtaactacgctatgcttggactactgggtcaaggaaacctcagtcacctctcc 711
Db 674 AGAGAT-----TATAGGCTTGACTACTGGGGCAAGGGACACAGGTACCGTCTCC 724
QY 712 tca 714
Db 725 TCA 727

RESULT 15
Q51540
ID Q51540 standard; DNA; 720 BP.
AC Q51540;
DT 23-MAY-1994 (first entry)
DE Coding sequence of polypeptide with affinity for fluorescein.
KW Monoclonal antibody; MAb; affinity; binding; antigen; diagnostics;
KW therapy; imaging; purification; biosensors; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_difference 619..621 /*tag= a
FT /*transl_except= ANT encodes Isoleucine.
FT misc_difference 658..660 /*tag= b
FT /*transl_except= GAA encodes Glycine.
PN US5260203-A.
PD 09-NOV-1993.
PF 02-SEP-1986: 902971.
PR 02-SEP-1986: US-902971.
PR 02-SEP-1987: US-092110.
PR 19-JAN-1989: US-299617.
PR 25-APR-1990: US-512910.
PA (ENZO-) ENZON INC.
PI Bird RE, Hardman K, Ladner RC;
DR F-PSDB; R43679.
PT Single chain poly:peptide for binding antigen - comprising light
PT and heavy chain antigen binding portions linked by peptide linker
PS Example 9; Figure 39; 78pp; English.
CC The single chain polypeptide is derived from the mature light and
CC heavy chains of a monoclonal antibody (MAb) and has affinity
CC for a given antigen (Fluorescein). It comprises a first
CC polypeptide comprising the antigen binding portion of of the light
CC chain variable region of an antibody and a second polypeptide
CC comprising the antigen binding portion of the heavy chain variable
CC region of an antibody and at least one peptide linker linking the
CC first and second polypeptide chains. The resulting single chain
CC polypeptide can be used in diagnostics, therapy
CC (in vivo and in vitro), imaging, purifications and biosensors.
CC This particular single chain binding molecule was designated
CC 18-2-3/TRY202' and contains one linker peptide.
SQ Sequence 720 BP; 180 A; 185 C; 175 G; 180 T;

```

Query Match 55.6%; Score 396.8; DB 1; Length 720;
 Best Local Similarity 76.0%; Pred. No. 2.6e-100;
 Matches 552; Conservative 0; Mismatches 147; Indels 27; Gaps 4;

```

QY 4 attgtatgacccagactcccaattcctgcttgcacagagagagaggttaccata 63
Db 7 AATGTGCTCACCAGCTCCAGCAATCATGCTGTCATCTCCAGGGGAAAGGTCCACATG 66
QY 64 acctcaagccagctcagactgtgagt---aatgatgtgcttggtaccacagaagcca 120
Db 67 ACCTGCGGGGCCAGCTCAAGTGTAAAGTTCAGTTACTTGCATCTGACAGGAGTCA 126
QY 121 gggcagctctccgaactcgtgatatactctgcacccaatcgctacactggagccctgat 180
Db 127 GGTGCTCTCCCAACACTCTGGGTTTATGGCACATCCAACTTGGCTTCTGGAGTCCCTGCT 186
QY 181 cgttcactggcagtgatggatggagcggatttcactttcaccatcagcactgtgcaggct 240

```

Search completed: November 30, 1999, 14:11:38
 Job time: 5223 sec

```

Db 187 CGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCACAATCAGCAGTGTGGAGCT 246
QY 241 gaagacctggcgagttatttctgtcagcagagattatag-----ctcgcctcgaggg 291
Db 247 GAAGATGCTGCCACATATTAATCTAGTGGCAGCAGTACAGTGGTTACCCACTCAGCTTCGGTGTCT 306
QY 292 gggaccaagctggaaaaataaaaggtggagcggttcagcgaggttggtctggcggtggc 351
Db 307 GGGACCAAGCTTGAGCTGGAGGTAAA-----TCTTCTGTTTCTGGTTCCGAA 354
QY 352 ggcgcaggttcaggtgaagagtcaggaacctgcccctgggtggcgccctcacagagcctg 411
Db 355 TCTAAATCTACTTACTGAAAGGAGTCAGGACCTGCTCTGGTGGCGCCCTCACAGAGCCTG 414
QY 412 tccatcacttgcaactgtctctgggttttcattaaaccaattatgggtgtacactgggttcgc 471
Db 415 TCCATCACTTGCACCGTCTCTGGGTTTTCATTAAACCACTATGGTGTACACTGGGTTCGC 474
QY 472 cagcctccaggaaaaggtctctgggtggagtgagtaatatgggtgggtgggaagcacaat 531
Db 475 CAGCCTCCAGGAAAGGCTCTGGAGTGGCTGGAGTAATATGGCTGTGTGGAACACAAAAT 534
QY 532 tataattgggtcttattgtccagactgagcatcagcaggaagacaactccaagagccaagt 591
Db 535 TATAATTCAGCTCTCATGTCCAGACTGAGCATCAAGAAAGACAATTCACAGAGCCAAAGTT 594
QY 592 ttcttaaaatgaacagctctgcaaacctgatgacacagccatgtactactgtgcca---gt 648
Db 595 TTCTTAAAAATGAACAGTCTGCAAAATGATGACACAGCCATATATACTACTGTGCCAAACGA 654
QY 649 cgggggggtaactacggtatgcttggactactgggttcagggaacctcagtcacctc 708
Db 655 CTGGAACGAACTCTTTTACTATGCTATGGACTATTGGGGTCAAGGAACCTCAGTCACCGTC 714
QY 709 tcttca 714
Db 715 TCCTAA 720

```


GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 14:00:09 ; Search time 383.27 Seconds
(without alignments)
4363.706 Million cell updates/sec

Title: US-08-940-544-4
Perfect score: 714
Sequence: 1 agtatgtgatgaccagac.....cctcagtcaccgtctctctca 714

Scoring table: IDENTITY_NUC

Searched: 3032314 seqs, 1171202697 residues

Database :

EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|----------|--------------------|
| 1 | 199.6 | 28.0 | 345 | 37 | AA10291 | AA10291 vt53a04.r |
| 2 | 190.6 | 26.7 | 471 | 43 | AI233978 | AI233978 EST230666 |
| 3 | 152 | 21.3 | 328 | 31 | AA295311 | AA295311 EST100471 |
| 4 | 146.2 | 20.5 | 370 | 31 | AA295093 | AA295093 EST100400 |
| 5 | 141.6 | 19.8 | 345 | 32 | AA335086 | AA335086 EST39457 |
| 6 | 139.2 | 19.5 | 413 | 31 | AA301347 | AA301347 EST14279 |
| 7 | 139 | 19.5 | 438 | 39 | AA880491 | AA880491 vx40h05.r |
| 8 | 138.6 | 19.4 | 393 | 20 | T27593 | T27593 EST100653.H |
| 9 | 136.4 | 19.1 | 463 | 22 | R69532 | R69532 vj82d09.rl |
| 10 | 136.4 | 19.1 | 352 | 31 | AA301361 | AA301361 EST14181 |
| 11 | 133.2 | 18.7 | 382 | 31 | AA295786 | AA295786 EST100987 |
| 12 | 130 | 18.2 | 349 | 32 | AA379044 | AA379044 EST91999 |
| 13 | 127.8 | 17.9 | 334 | 31 | AA327254 | AA327254 EST30547 |
| 14 | 126.8 | 17.8 | 357 | 32 | AA361497 | AA361497 EST71040 |
| 15 | 126.4 | 17.7 | 303 | 31 | AA300788 | AA300788 EST13648 |
| 16 | 125.8 | 17.6 | 516 | 22 | R67559 | R67559 yi42h11.rl |
| 17 | 125 | 17.5 | 303 | 31 | AA300891 | AA300891 EST14031 |
| 18 | 123 | 17.2 | 381 | 20 | T29114 | T29114 EST69430.Hu |
| 19 | 122.8 | 17.2 | 364 | 28 | AA098196 | AA098196 mm86h06.r |
| 20 | 122.2 | 17.1 | 363 | 32 | AA367405 | AA367405 EST78511 |
| 21 | 122.2 | 17.1 | 432 | 35 | AA544384 | AA544384 vk33f04.r |
| 22 | 121.4 | 17.0 | 517 | 63 | AI989649 | AI989649 ws26f12 |
| 23 | 120.8 | 16.9 | 823 | 29 | AA170256 | AA170256 ms87g10 |
| 24 | 119.8 | 16.8 | 364 | 20 | T27579 | T27579 EST100049 |
| 25 | 119.8 | 16.8 | 396 | 37 | AA691311 | AA691311 vs14f01.r |
| 26 | 119.2 | 16.7 | 363 | 31 | AA300651 | AA300651 EST13764 |
| 27 | 119 | 16.7 | 301 | 32 | AA377295 | AA377295 EST98837 |
| 28 | 118.4 | 16.6 | 400 | 34 | AA456778 | AA456778 zw27f11.r |
| 29 | 117.6 | 16.5 | 403 | 32 | AA345486 | AA345486 EST51505 |
| 30 | 117.4 | 16.4 | 395 | 20 | T29112 | T29112 EST69384.Hu |
| 31 | 117 | 16.4 | 511 | 34 | AA464794 | AA464794 zx83h07.r |
| 32 | 116.4 | 16.3 | 442 | 33 | AA405415 | AA405415 zu56d02.r |
| 33 | 116.4 | 16.3 | 425 | 51 | AI734035 | AI734035 zu56d02.y |
| 34 | 115.8 | 16.2 | 332 | 20 | T29342 | T29342 EST77181.Hu |
| 35 | 114.8 | 16.1 | 335 | 31 | AA318377 | AA318377 EST20620 |
| 36 | 114.8 | 16.1 | 405 | 32 | AA360223 | AA360223 EST69341 |
| 37 | 114.8 | 16.1 | 363 | 34 | AA464313 | AA464313 zx78c12.r |
| 38 | 114.6 | 16.1 | 346 | 31 | AA300582 | AA300582 EST13427 |
| 39 | 113.4 | 15.9 | 282 | 31 | AA300491 | AA300491 EST13404 |
| 40 | 111.6 | 15.6 | 503 | 22 | R28232 | R28232 yhs7e08.rl |
| 41 | 111.2 | 15.6 | 398 | 33 | AA423447 | AA423447 ve80a03.r |
| 42 | 110.8 | 15.5 | 292 | 31 | AA295508 | AA295508 EST100751 |
| 43 | 110.2 | 15.4 | 418 | 34 | AA515239 | AA515239 dg69c07.s |
| 44 | 109.6 | 15.4 | 469 | 49 | AI645111 | AI645111 ms87g10.y |
| 45 | 109.4 | 15.3 | 424 | 41 | AI045896 | AI045896 UI-R-CO-j |

ALIGNMENTS

| RESULT | 1 | |
|------------|---|------------|
| LOCUS | AA710291 | |
| DEFINITION | AA710291 345 bp mRNA EST 24-DEC-1997 cl53404.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus CDNA clone IMAGE:1166766 5' similar to gb:X02484.rnal IG KAPPA CHAIN PRECURSOR V-IV REGION (HUMAN); gb:J00560 mouse ig kappa mrna from moe231 & other myeloma mrna 3' (MOUSE);, mRNA sequence. | |
| ACCESSION | AA710291 | |
| NID | g2720209 | |
| VERSION | AA710291.1 | GI:2720209 |
| KEYWORDS | EST. | |
| SOURCE | house mouse. | |
| ORGANISM | Mus musculus | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| AUTHORS | 1 (bases 1 to 345) Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R. | |
| TITLE | The WashU-HMMI Mouse EST Project | |
| JOURNAL | Unpublished (1996) | |
| COMMENT | On Sep 12, 1996 this sequence version replaced gi:1326564. | |

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28mi3 rev2 ET from Amersham.
High quality sequence stop: 140.

| | |
|----------|--|
| FEATURES | |
| SOURCE | |

```

1. 343
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/map="961D03; 16; 14; 21q"
/clone="IMAGE:1166766"
/clone.lib="Barstead mouse irradiated colon MPLRE7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'GCTTACGATCTGAGTGGAGCGCCGCTTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGGATCCCTTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library constructed by Bob Barstead."
95 a 81 c 88 g 81 t

```

```

      95 a      81 c      88 g      81 t
      vccol: library constructed by bob parstead.

      Query Match      28.0%      Score 199.6;      DB 37;      Length 345;
      Best Local Similarity      81.4%;      Pred. No. 1.8e-47;
      Matches 259;      Conservative      0;      Mismatches 49;      Indels 10;      Gaps 2;
      Mismatches 49;      Indels 10;      Gaps 2;

```

[illegible]

| | | | |
|----|-----|---|-----|
| Db | 88 | ACCTGCAA--CCAGTCAGAAATGTGGGTACTAATGTAGCTGGTATCAAGAGAAACACAGGG | 146 |
| Qy | 124 | cagtcctccgaactgctgatactactctgcatccaactcactggagtcctcgatcgc | 183 |
| Db | 147 | CAATCTCCTAGACACTGATTTACTCGGCATCCTACCGGTACAGTGGAGTCCCTGATCGC | 206 |
| Qy | 184 | ttcaactgcagctggatatggcagcgatttcactttccactcagcactdgcagcgctgaa | 243 |
| Db | 207 | TTCACAGGCAGTGGATCTGGGACAGATTTTCACTCTCACCATCAGCAATGTGACGCTTGAA | 266 |
| Qy | 244 | gacctggcagctttattctctgcagcaggattatagct-----cgctccgagggggg | 294 |
| Db | 267 | GACTTGGCAGAGTATTTCTGCTAGCAATATACAGGCTTTCGCTACACGTTTCGAGGGGGG | 326 |
| Qy | 295 | accaagctggaaataaaa | 312 |
| Db | 327 | ACCAAGCTGGAAATAAAA | 344 |

| RESULT | 2 | | | | | | |
|------------|---|------------|----------------------|--------------|------------|-------------|--|
| AI233978/C | | | | | | | |
| LOCUS | | AI233978 | 471 bp | mRNA | EST | 31-JAN-1999 | |
| DEFINITION | | EST230656 | Normalized rat lung, | Bento Soares | Rattus sp. | cDNA clone | |
| | | RL0CT66_3' | end, mRNA sequence. | | | | |

accession
 NID 93817858
 version AI233978.1
 keywords GR:3817858
 source EST.
 organism Rattus sp.
 Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 reference 1 (bases 1 to 471)
 authors Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
 Kerlavage,A.R. and Adams,M.D.
 title Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
 Gene Index
 journal Unpublished (1998)
 comment On Apr 18, 1995 this sequence version replaced qi:775307.

Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

```

seq primer: nt3 21.
location/Qualifiers
  1..471
    /organism="Rattus sp."
    /db_xref="ATCC (inhost):2038523"
    /db_xref="taxon:10118"
    /map="17"
    /clone="RLUCT66"
    /clone_lib="Normalized rat lung, Bento Soares"
    /note="Organ: lung; Vector: pT73Pac; Site_1: EcoRI;
    Site_2: NotI"
BASE COUNT      129 a      94 c      132 g      116 t
ORIGIN

```

| | Query Match | 26.7%; | Score 190.6; | DB 43; | Length 471; |
|----|-----------------------|--|--------------------|-----------|-------------|
| | Best Local Similarity | 80.5%; | Pred. No. 8.1e-45; | | |
| | Matches 223; | Conservative 0; | Mismatches 54; | Indels 0; | Gaps |
| Qy | 4 | atttgatgaccacagactcccaaatctctgttgcattcagcaggagacaggggttaccata | 63 | | |
| | | | | | |
| | | | | | |
| | | | | | |
| Db | 443 | ATTGTGATGACTCAGTCTCCCATCCATGTCCATATCATGTAGGACACAGGGTCAACATG | 384 | | |
| Qy | 64 | acctgcaagcccgatgcagagtgtgagtaaatgatgtggcttggtaccacaacagaagccaggg | 123 | | |
| | | | | | |
| | | | | | |
| Db | 383 | AACTGCAAGGCCAGTCAAAATGTGGTCTTACTGTACTGTGTACCAACAGAAAAACAGGG | 324 | | |

| | | | |
|------------|-----|--|-------------|
| QY | 124 | cagctccgaactcgtgatactactctgcacccaatcgctacatcgtagtcctcgtatcgc | 183 |
| Db | 323 | CAGCTCCTAAACTGCTTTATCTTCAAGACCTCAACACGGGCACGGGAGTCCTCGTGATCGC | 264 |
| QY | 184 | ttcacctggcaggatattggcagcatcttcaactttaccatcacacactgtgcagctgaa | 243 |
| Db | 263 | TTCACAGGAGATGATCTGACACCATTTCACTTTCCACCATCACACAATGCAGGCTGA | 204 |
| QY | 244 | gacctggcagtattattcttgtcgcagcaggattatgct | 280 |
| Db | 203 | GACCTGGCTGTTTTATTACTGTGATGTCAGTCTAATCCT | 167 |
| RESULT | 3 | | |
| AA295311 | | | |
| LOCUS | | 328 bp mRNA EST | 18-APR-1997 |
| DEFINITION | | EST100471 Pancreas tumor I Homo sapiens CDNA 5' end similar to immunoglobulin kappa light chain, VJ regions, mRNA sequence. | |
| ACCESSION | | AA295311 | |
| NID | | 91947646 | |
| VERSION | | AA295311.1 | GI:1947646 |
| KEYWORDS | | EST. | |
| SOURCE | | human. | |
| ORGANISM | | Homo sapiens | |
| REFERENCE | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| AUTHORS | | 1 (bases 1 to 328) | |
| | | Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Van-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glocke,A., Gnehm,C.I., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Klotz,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligriano,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,I.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D., Lu,K., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haselaine,W.A., Fields,C., Fraser,C.M. and Venter,J.C. | |
| TITLE | | Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence | |
| JOURNAL | | Nature 377 (6547 Suppl), 3-174 (1995), | |
| MEDLINE | | 96026280 | |
| COMMENT | | On Sep 1, 1995 this sequence version replaced. Other_ESTs: HCl619106 Bioinformatics The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850 USA Tel: 3018699056 Fax: 3018699423 Email: arkerlavet@igr.org For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html) Seq primer: M13 Reverse Location/Qualifiers 1..328 /organism="Homo sapiens" /db_xref="ATCC (inhost):190766" /db_xref="taxon:9606" /clone_lib="Pancreas tumor I" /dev_stage="adult" | |
| FEATURES | | source | |
| BASE COUNT | | 73 a 93 c 77 t 7 others | |
| ORIGIN | | Ecoki; Site_2: XhoI" | |

| DEFINITION | ACCESSION NID | VERSION | KEYWORDS | SOURCE ORGANISM | REFERENCE AUTHORS | TITLE JOURNAL COMMENT |
|------------|------------------|---------|----------|--------------------|----------------------|-----------------------------|
|------------|------------------|---------|----------|--------------------|----------------------|-----------------------------|

COMMENT

FEATURES
SOURCE

| BASE COUNT ORIGIN | Query Match Best Local Matches | |
|----------------------|--------------------------------------|---|
| QY | 336 | 1 |
| Db | 58 | 1 |
| QY | 396 | 9 |
| Db | 118 | 6 |
| QY | 454 | - |
| Db | 178 | 1 |

| | |
|----|-------|
| QY | 510 a |
| Dd | 238 T |

| DEFINITION | ACCESSION NID | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|------------|------------------|---------|----------|--------|----------|-----------|---------|-------|---------|---------|
|------------|------------------|---------|----------|--------|----------|-----------|---------|-------|---------|---------|

| FEATURES | SOURCE |
|------------|--------|
| BASE COUNT | |

| BASE COUNT ORIGIN | Query Match Best Local Matches |
|----------------------|--------------------------------------|
| QY | 336 |
| Db | 58 |
| QY | 396 |
| Db | 118 |
| QY | 454 |
| Db | 178 |
| QY | 510 |
| Db | 238 |


```

/clone_lib="T-cell lymphoma"
/cell_type="T-lymphocyte"
/Note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      88 a   96 c   84 g   87 t   2 others
ORIGIN

Query Match      17.8%; Score 126.8; DB 32; Length 357;
Best Local Similarity 70.7%; Pred. No. 1.9e-26;
Matches 195; Conservative 0; Mismatches 79; Indels 2; Gaps 2;

QY 4 attgtgatgaccagactcccaaatctctgtgtatcagcagagacagggttaccata 63
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 69 ATCCAGATGACCCAGTCTCCATCCCTCTGTGTGATCTAGGACAGAGTCCACATC 128
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 acctcaaggccagcagtgatgagtaagtgtggtgtgtaccacaagaagccagg 123
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 129 ACTTCCCGGGCAAGNAGAGAAATTAACAACTATTAAATTGGTATCAACAGAACCCAGG 188
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 124 cagctccgaactgctgatatactctgtcatcccaatc-gctacacgtgagtcctgtatcg 182
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 189 AAGGCCCTAAGCTCTGATACGCTACGGTCATCCAGTCTGCAAGTGGGGTCCCATCAAG 248
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 ctctcagcagtgatgataggagcga-ttcaactttccaccatcagcactgtcaggctg 241
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 249 GTTCAGTGGCAGTGATCTGGGACAGATTTTCACTCTCANCATCATCGCGTCTGCAACCTG 308
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 aagacctgagcttattcttctgcagcaggattata 277
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 309 AAGATTTTGCAACTTACTACTGTCACACAGAGTTACA 344
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AA300788
LOCUS      AA300788      303 bp      mRNA      EST      18-APR-1997
DEFINITION EST13648 Testis tumor Homo sapiens cDNA 5' end similar to
            immunoglobulin kappa light chain, V region, mRNA sequence.
ACCESSION  AA300788
NID        91953120
VERSION    AA300788.1 GI:1953120
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 303)
AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,D.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
TITLE      Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
JOURNAL    Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE    96026280
COMMENT    On Sep 12, 1996 this sequence version replaced gi:1397991.
            Other ESTs: THC167177
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research

```

9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavetigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

source
 1..303
 /organism="Homo sapiens"
 /db_xref="AYCC (inhost):192224"
 /db_xref="taxon:9606"
 /map="5q32-33"
 /clone_lib="Testis tumor"
 /sex="male"
 /dev_stage="adult"
 /note="Organ: testis; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

BASE COUNT 69 a 88 c 72 g 72 t 2 others
 ORIGIN

Query Match 17.7%; Score 126.4; DB 31; Length 303;
 Best Local Similarity 71.2%; Pred. No. 2.3e-26;
 Matches 166; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 4 attgtgatgaccagactcccaaatctctgtgtatcagcagagacagggttaccata 63
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 71 ATCCAGATGACCCAGTCTCCATCCCTCTGTGTGATCTAGGACAGAGTCCACATC 130
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 64 acctcaaggccagcagtgatgagtaagtgtggtgtgtaccacaagaagccagg 123
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 131 ACTTCCCGGGCAAGTCCAGCATTAGCACCTATTAAATTGGTATCAGCAGAACCCAGG 190
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 124 cagctccgaactgctgatatactctgtcatcccaatc-gctacacgtgagtcctgtatcg 183
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 191 AAGCCCTAATCTCTGTGATCTCTGTGATCTGTCATCCAGTTTGCAAGTGGGTCCCATCAAG 250
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 184 ttcactggcagtgatgataggacgatttcacatttcaccatcagcactgtgca 236
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 251 TTCAGTGTCACTGGTCTGGGACAGATTTCACCTCTCACCATCAGACGTCTGTGCA 303

Search completed: November 30, 1999, 14:00:11
 Job time: 4539 sec

```

Query Match      55.5%;   Score 396.2; DB 1;   Length 738;
Best Local Similarity 75.5%;   Pred. No. 6.5e-110;
Matches 355; Conservative 0; Mismatches 153; Indels 27; Gaps 4;

y      4 attgtgtagaccagactcccaaatctcgttgttatcagcaggagacaggttaccata 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b      7 ATCTCTGCTGACACAGTCTCCAAAATTCCTGCTGTATCAGCAGGAGACAGGTTACCAT 66

```


RESULT 5
US-08-634-783A-3
; Sequence 3, Application US/08634783A
; Patent No. 3861276
; GENERAL INFORMATION:
; APPLICANT: KWAK, JU-WON
; APPLICANT: HAN, MOON-HI
; APPLICANT: CHOI, BYUNG-KWON
; TITLE OF INVENTION: C-DNAS ENCOD

TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 PAGE MILL ROAD
 CITY: PALO ALTO
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/752,844
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Schilff, J. Michael
 REGISTRATION NUMBER: 40,253
 REFERENCE/DOCKET NUMBER: 30414-20002.21
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 304 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PS-08-752-844-32

| | | | | |
|-----------------------|--------------|-------------------|--------------|------------|
| Query Match | 39.6% | Score 282.8 | DB 4 | Length 304 |
| Best Local Similarity | 97.6% | Pred. No. 4.7e-76 | | |
| Matches 287 | Conservative | 0 | Mismatches 7 | Indels 0 |
| | | | | Gaps 0 |

| | | | |
|----|-----|---|-----|
| QY | 358 | caagtcgacgtgaaggagtcagagacctggcctggtggcgccctcacagagacctgtccatc | 417 |
| | | | |
| Db | 1 | CAGGTGCAGCTGAAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTGTCCATC | 60 |
| | | | |
| QY | 418 | acttgcaactgtctctggggttttcattaaaccaattatggtgtacactgggttcgccagcct | 477 |
| | | | |
| Db | 61 | ACTTGCACCTCTCTGGGTTTTTCATTAACCAAGCATATGTTGTACACTGGTTCGCCAGCCT | 120 |
| | | | |
| QY | 478 | ccgggaaaagggtctgagtcggctgggagtgtaatatggcctggtgggaagcacaaattataat | 537 |
| | | | |
| Db | 121 | CCAGGAAAGGGTCTGGAGTGGCTGGGAGTAATATGGGCTGGTGGGAAGCACAAATATAAT | 180 |
| | | | |
| QY | 538 | tcggctcttatgtccagactgagcatcagcaagagacaactcccaagagccaagttttctta | 597 |
| | | | |
| Db | 181 | TCGGCTCTCATGTCCAGACTTGACATCAGCATCAGCAAGAGACAATCCCAAGAGCCCAAGTTTCTTA | 240 |
| | | | |
| QY | 598 | aaatgaacagtcgtcaaacctgatgacacagccatgtactactgtgccagtcgg | 651 |
| | | | |
| Db | 241 | AAAATGAACAGTCTGCAAACTGATGACACAGCCATGTACTACTGTGCCAGAGGG | 294 |
| | | | |

| | |
|--|---|
| RESULT | 7 |
| US-08-752-844-3 | |
| ; Sequence 3, Application US/08752844 | |
| ; Patent No. 5935821 | |
| ; GENERAL INFORMATION: | |
| ; APPLICANT: Chatterjee, Malaya | |
| ; APPLICANT: Foon, Kenneth A. | |
| ; APPLICANT: Chatterjee, Sunil K. | |
| ; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE | |
| ; TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA | |
| ; NUMBER OF SEQUENCES: 66 | |
| ; CORRESPONDENCE ADDRESS: | |

Sequence 3, Application US/08/52844
Patent No. 5935821
GENERAL INFORMATION:
APPLICANT: Chatterjee, Malaya
APPLICANT: Foon, Kenneth A.
APPLICANT: Chatterjee, Sunil K.
TITLE OF INVENTION: MONOCLONAL ANTIBODY I7A7 AND USE FOR THE
TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:

```

ADDRESS: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752-844-3
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/POCKET NUMBER: 30414
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..456
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 58
US-08-752-844-3

```

```
Query Match      39.5%; Score 282; DB 4; Length 458;
Best Local Similarity 86.4%; Pred. No. 9.8e-76;
Matches 325; Conservative 0; Mismatches 45; Indels
```

| | | | |
|-----|----|--|-----|
| 339 | Qy | ctctggcggctggcggatcgcagatgcaggtgaaggagtcaggacctggcctggctggcgcc | 398 |
| | | | |
| 39 | Db | CCCAAGCTGTCTGCTGTCACAGTGCAGGTGAAGAGTCAGGACCTTCTTGGTGGCCCC | 98 |
| | | | |
| 399 | Qy | ctcacagagcctgcataactgcactgtctctgggttttcaataaccaaattatgggtc | 458 |
| | | | |
| 99 | Db | CTCACAGAGGCTGTCCATCACAATGCACCTGTCTCAGGGTTCTCATTAACCACTATATGGT | 158 |
| | | | |
| 459 | Qy | acactgggttcgcagcctccaggaagggtctggagtcgggtgggagtaatatgggctgg | 518 |
| | | | |
| 159 | Db | AAGCTGGATTGCCAGCCTCCACGAAAGGGTCTGGAGTGGCTGGGAGCAATTTGGGGTCA | 218 |
| | | | |
| 519 | Qy | tggaagcacaaaataattcggctcttatgtccagactgagcatcagcaaggacaactc | 578 |
| | | | |
| 219 | Db | CGGACCCACAAATATCATTTCAGCTCTCATATCCAGACTGAGCATCAGCAAGGATAACTC | 278 |
| | | | |
| 579 | Qy | caagagcccaagtttcttaaaaatgaacagctctgcacaaactgatgcacagccatgtacta | 638 |
| | | | |
| 279 | Db | CAAGAGCCAAAGTTTCTTAATAACTGAACAGCTGTGCAAACTGATGCACCGGCCACGTACTA | 338 |
| | | | |
| 639 | Qy | ctgtggcagctcgggggggtaactcacgctatgctttggactactgggttcagggaacctc | 698 |
| | | | |
| 339 | Db | CTGTGCCAAACTGGG-----TAACCTACGATGCTCTGGACTACTGGGGTCAAGGAACCTC | 392 |
| | | | |
| 699 | Qy | agtcaccgtctcctca | 714 |
| | | | |
| 393 | Db | AGTCACCGTCTCCTCA | 408 |
| | | | |

RESULT 8

US-08-752-844-37
 : Sequence 37, Application US/08752844
 : Patent No. 5935821
 : GENERAL INFORMATION:
 : APPLICANT: Chatterjee, Malaya
 : APPLICANT: Foon, Kenneth A.
 : APPLICANT: Chatterjee, Sunil K.
 : TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
 : TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
 : NUMBER OF SEQUENCES: 66
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: MORRISON & FOERSTER
 : STREET: 755 PAGE MILL ROAD
 : CITY: PALO ALTO
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304-1018
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/752,844
 : FILING DATE:
 : CLASSIFICATION: 424
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Schiff, J. Michael
 : REGISTRATION NUMBER: 40,253
 : REFERENCE/DOCKET NUMBER: 30414-20002.21
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 813-5600
 : TELEFAX: (415) 494-0792
 : TELEX: 706141
 : INFORMATION FOR SEQ ID NO: 37:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 291 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : US-08-752-844-37

Query Match 39.5%; Score 282; DB 4; Length 291;
Best Local Similarity 98.3%; Pred. No. 8e-76;
Matches 285; Conservative 0; Mismatches 5; Indels

| | | | |
|----|-----|---|-----|
| Qy | 358 | caggctgcagctgaagdagtcagacactgctggcgtccctcacagagcctgtccatc | 417 |
| | | | |
| | | | |
| Db | 1 | CAGGTGCAGCTGAAGGAGTCAGACACTGGCCTGGTGGCGCCCTCACAGAGCCTGTCCATC | 60 |
| | | | |
| Qy | 418 | acttgcactctctcggttttcatlaaccaatatgggtacactgggttcgcagcct | 477 |
| | | | |
| | | | |
| Db | 61 | ACTTGCACTGCTCTGGTTTTTCATTAAACCAGCTATGGTGTACACTGGGTTCGCCAGCCT | 120 |
| | | | |
| Qy | 478 | ccaggaaagggctcgagtggctgggagtaatatggcctgggtggaagcacaattataat | 537 |
| | | | |
| | | | |
| Db | 121 | CCAGGAAAAGGGCTGTGAGTGGCTGGGAGTAATAATGGGCTGGTGGAGAGCACAAATTATAAT | 180 |
| | | | |
| Qy | 538 | tcggctcttatgtccagactgaacatcagcaaggaacaactccaagaaccaagtttttcta | 597 |
| | | | |
| Db | 181 | TCGGCTCTCATGTCCAGACTGACATCAGCAAGAACAATCCAAGAGCCAAAGTTTTCCTTA | 240 |
| | | | |
| Qy | 598 | aaaatgaacagctgtcaaacatgatgacacagccatgtactactgtgccag | 647 |
| | | | |
| Db | 241 | AAAATGAACAGTCTGCAAACTGATGACACAGCCATGCTACTGTGCCAAG | 290 |
| | | | |

RESULT 9
US-08-752-844-65
; Sequence 65, Application US/08752844
; Patent No. 5935821

Qy 292 gggaccaagctggaat 308
 |||||
 Db 301 GGGACCAAGCTGGAGAT. 317

RESULT 14

```

US-08-463-298-2
/ Sequence 2, Application US/08463298
/ Patent No. 5824304
/ GENERAL INFORMATION:
/ APPLICANT: Papayannopoulou, Thalia (USA only)
/ APPLICANT: Board of Regents, U
/ APPLICANT: Washington (except USA)
/ TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
/ TITLE OF INVENTION: CELLS
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: C/O FISH & NEAVE
/ STREET: 1251 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10020
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/463,298
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/11060
/ FILING DATE: 11-NOV-1993
/ APPLICATION NUMBER: US 07/977,702
/ FILING DATE: 13-NOV-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haley Jr., James F.
/ REGISTRATION NUMBER: 27,794
/ REFERENCE/DOCKET NUMBER: B173CIP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 596-9000
/ TELEFAX: (212) 596-9090
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 318 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ US-08-463-298-2

```

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 38.7% | Score 276.2; | DB 3; | Length 318; |
| Best Local Similarity | 94.6% | Pred. No. 4.6e-74; | | |
| Matches 300: Conservative | 0; | Mismatches 8; | Indels 9; | Gaps 1; |

| | | | |
|----|-----|--|-----|
| QY | 1 | agtatgtgatgaccagactcccaaatctctgtctgtatcagcaggagacagggttacc | 60 |
| Db | 1 | AGTATTGTGATGATCCAGACTCCCAAAATCCTGTTGTTTACAGAGGACAGGGTTACC | 60 |
| QY | 61 | ataacctcaaggccagtcagagcttgagtaatgatgtggcttggttaccaacagaagcca | 120 |
| Db | 61 | ATAACTGTCAAGGCCAGTCAGAGTGTGACTAATGATGTAGCTTGGTACCAAGAGGCCA | 120 |
| QY | 121 | gggcagctctccgaactgctgtatactctgcatccaatcgctacactggagttccctgat | 180 |
| Db | 121 | GGSCAGTCTCCTAACTGCTGATATATTATGCATCCAAATCGTACACTGGAGTCCCTGAT | 180 |
| QY | 181 | cgcttcaactggcagtggaatatgggacggatttcaactttcaccatcagcactgucaggct | 240 |
| Db | 181 | CGCTTCACGTGGCAGTGGATGGACGGATTTTCATTTCCACATCAGCACTGTGAGGCT | 240 |

| | | | |
|----|-----|---|-----|
| Qy | 241 | gaagacctggcagttattttctgcacgagattatagct-----cactcgagggg | 291 |
| | | | |
| Db | 241 | gaagacctggcagttattttctgcacgagattatagctctccgtacacgttcggagg | 300 |
| | | | |
| Qy | 292 | gggaccaagctgggaaat | 308 |
| | | | |
| Db | 301 | gggaccaagctgggaaat | 317 |
| | | | |

RESULT 15

```

US-08-436-339A-2
; Sequence 2, Application US/08436339A
; Patent No. 5843438
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 03-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173C1P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-436-339A-2

```

Query Match 38.7%; Score 276.2; DB 3; Length 318;
Best Local Similarity 94.6%; Pred. No. 4.6e-74;
Matches 300; Conservative 0; Mismatches 8; Indels 9; Gaps 1;

| | | | |
|----|-----|---|-----|
| Qy | 1 | agtatgtgataccagagatcccaattcctgctgtatcagcagagacagaggttacc | 60 |
| Db | 1 | AGTATTGTGATGACCCAGATCCCAAATCCGTGTTTCAGCAGGAGACAGGTTACC | 60 |
| Qy | 61 | ataacctcaaggccagctcagagctgagtaatgatgcttggtaccacaagaacca | 120 |
| Db | 61 | ATAACCTCAGGCCAGTCAGATGTGACTAATGATGTAGCTTGGTACCAACAGAAGCCA | 120 |
| Qy | 121 | gggcagctccgaactgctgatatctctgcatccaatcgctacactggagtcctcgat | 180 |

Db 121 GGGCAGTCTCCTAACTGCTGATATATATGTCATCCAAATCGCTACACTGGAGTCCCTGAT 180
Qy 181 cgcttcaactggcagtgatattgggacgagatttcactttcaccatcagcactgtgcaggct 240
|||||
Db 181 CGCTTCACTGGCAGTGGGATATGGGACGGATTTCACCTTCCACATCAGCACTGTGCAGGCT 240
Qy 241 gaagacctggcagttattttctgtcagcaggattatagct-----cgctcggagg 291
|||||
Db 241 GAAGACCTGGCAGTTTATTCTGTGACAGGATTATAGCTCTCCGTACACGTTGCGAGG 300
Qy 292 gggaccaagctggaaat 308
|||||
Db 301 GGGACCAAGCTGGAGAT 317

Search completed: November 30, 1999, 14:10:21
Job time: 5147 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 14:09:30 ; Search time 545.98 Seconds
(without alignments)
5591.774 Million cell updates/sec

Title: US-08-940-544-4

Perfect score: 714

Sequence: 1 agtattgtgagccagac.....cctcagtcacogtctctca 714

Scoring table: IDENTITY_NUC

Searched: 780561 seqs, 2137953050 residues

Database : GenEmbl.*

1: gb_bal.*
2: gb_baz.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vl.*
17: em_fun.*
18: em_hcg.*
19: em_humi.*
20: em_hum2.*
21: em_in.*
22: em_om.*
23: em_or.*
24: em_ov.*
25: em_pat.*
26: em_ph.*
27: em_pl.*
28: em_ro.*
29: em_sts.*
30: em_sy.*
31: em_un.*
32: em_vl.*
33: gb_btg1.*
34: gb_btg2.*
35: gb_in1.*
36: gb_in2.*
37: em_bal.*
38: em_ba2.*
39: em_hum3.*
40: em_hum4.*
41: gb_pr4.*
42: gb_btg3.*
43: gb_btg4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match Length | ID | Description |
|------------|-------------|--------------|----|-------------|
| ----- | | | | |

| RESULT | 1 | MUSVKLI | MUSVKLI | 738 bp | DNA | ROD | 14-DEC-1995 |
|------------|---|---------|---------|--------|-----|-----|-------------|
| LOCUS | | | | | | | |
| DEFINITION | | | | | | | |
| ACCESSION | | | | | | | |
| NID | | | | | | | |
| VERSION | | | | | | | |
| KEYWORDS | | | | | | | |
| SOURCE | | | | | | | |
| ORGANISM | | | | | | | |
| REFERENCE | | | | | | | |
| AUTHORS | | | | | | | |
| TITLE | | | | | | | |
| JOURNAL | | | | | | | |
| REFERENCE | | | | | | | |
| AUTHORS | | | | | | | |

ALIGNMENTS

| | | | | |
|------------|--------|-----|-----|-------------|
| MUSVKLI | 738 bp | DNA | ROD | 14-DEC-1995 |
| LOCUS | | | | |
| DEFINITION | | | | |
| ACCESSION | | | | |
| NID | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |


```

FEATURES             Location/Qualifiers
     source            1..840
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /note="recombinant gene"
     CDS               1..840
                        /codon_start=1
                        /product="IgM/kappa antibody"
                        /protein_id="CAB08854.1"
                        /db_xref="PID:e1167934"
                        /db_xref="PID:g2108309"
                        /db_xref="GI:2108309"
                        /translation="MAQVGDIDQMTQSPPSLAYSGEKVTMSCKSSQSVLYSSNQKNY
LAWQKPGSPKLLIYWAISRESGVDRFTGSGDFTLTISVQAEADLAVYICHQ
YLSLFTFGAGTKLEIKREGSKSGSEKLESEVNLVSGPGLVQPSLSITCVSG
FSLTSYGVHWVQSGKGLWGLVIRWGGSTDYNAAFMRSLITKDNKSKQVFFRMS
LQADDTAIYCAKKNYGYNDYWGQGTSTVTSAPKTTPGAAGKLLISEDLNDIK
DEL"
     misc_feature      19..354
                        /note="Light chain variable domain (VL) coding sequence"
     misc_feature      355..402
                        /note="Linker peptide coding sequence"
     misc_feature      403..759
                        /note="Heavy chain variable domain (Vh) coding sequence"
     misc_feature      760..774
                        /note="Heavy chain first constant domain (CH1) fragment
coding sequence"
     misc_feature      787..819
                        /note="c-myc tag coding sequence"
     misc_feature      826..837
                        /note="KDEL coding sequence"
     BASE COUNT       230 a 210 c 209 g 191 t
     ORIGIN
1 atgtgtgaccagactcccaattctgtgtatcagcagagacagagttaccata 63
2  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 22 ATCCAGATGACAGAGTCTCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 81
4  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5 64 acctgcaaggccagctcagagtggt-----gagtaaatgatgtggtctgg 105
6  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 82 AGCTGTAGTCCAGTCAAGAGTGTATATACAGTTCAAAATCAGAGAAGTACTT 141
8  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9 106 taccacagaagcaggcagctccgaactgctgatatatactgtcatccaatcgctac 165
10 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
11 142 TACCAGCAGAAACAGGCGAGTCCTCTAAACTGCTGATCTACTGGGCTTCCACTAGGAA 201
12 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13 166 actgagtcctgctgcttcaactgagtgatgtaggacggattcaacttaccatc 225
14 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15 202 TCCGGTGTCTGATCGCTTCCAGGAGTGATGATCTGGACAGATTTTACCTTACCATC 261
16 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 226 agcactgtcaggctgaagacgttcgagttatttctgtcagcaggattatagctcgc-- 284
18 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 262 AGCAGTGTACAAGCTGAAGACTCGCAGTTTATTACTGTATCAATACCTCTCCCTCGCTC 321
20 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 284 ----tcggaggggggacacagctggaataaaagggtggagcggttcacggcgaggtggc 339
22 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 322 ACGTTCGGTGGGGACCAAGCTGGAATAAAACGGGAGGTAATCTCAGGATCTGGC 381
24 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 340 tctg---gcgggtgcgagtcaggtcaggctgaagagtcaggacacctgctggtggcg 396
26 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 382 TCCGAATCCAAACTCGAGTCTGAGGTGAACCTGTTGGAATCTGACCTGGCCTAGTGCAG 441
28 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
29 397 ccctcacagagcctgtccatcactgtcactgtctctggttttccattaaacaaattatggt 456
30 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 442 CCCTCACAGAGCCCTGTCCATAACCTGCACAGCTCTCTGGTTCTCATTAACCTAGCTATGGT 501
32 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 457 gtacactgggttcgccagcctccaggaaagggtctgagtggtgggagtaatatgggct 516
34 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 502 GTACACTGGGTCCGCCAGTCTCCAGGAAGGGTCTCGAGTGGCTGGAGTGATATCGAGA 561
 Qy 517 ggtggaagcacaaattataattcggctcttattgtccagactgagcatcagcaggacaac 576
 Db 562 GGTGGAAGCACACACTACAATGCAGCTTTCATGTCCAGACTGAGCATCACCAGGACAAC 621
 Qy 577 tcaagagccaagttttcttaaaaaatgaacagctctcaaaactgatgacacagccatgtac 636
 Db 622 TCCAAGAGCCAAGTTTCTTTAAATGAACAGTCTCAAGCTGATGACACTGCCATATAC 681
 Qy 637 tactgtgcagctcggggggtaactcagcgtctatgttgactactgggtcagaagaac 696
 Db 682 TACTGTGCCAAAATTAATACTAGTACAGGGGCTATGACTACTGGGTCAAGGAACC 741
 Qy 697 tcagtcacgcgtctcctca 714
 Db 742 TCAGTCACAGTCTCTGCA 759

RESULT 3
 LOCUS SYN7A41
 DEFINITION Synthetic single-chain antigen-binding protein gene (7A4-1/212 SCA) constructed from the antigen-binding (Fab) fragment of mouse catalytic monoclonal antibody NPN43C9 and linker DNA, partial cds.
 ACCESSION M68968
 NID 9207839
 VERSION M68968.1
 KEYWORDS 7A41/212 SCA protein; fusion protein; monoclonal antibody; single-chain antigen-binding protein.
 SOURCE Mouse DNA and synthetic DNA.
 ORGANISM artificial construct.
 REFERENCE 1 (bases 1 to 743)
 AUTHORS Gibbs, R.A., Posner, B.A., Filpula, D.R., Dodd, S.W., Finkelman, M.A., J., Lee, T.K., Wroble, M., Whitlow, M. and Benkovic, S.J.
 TITLE Construction and characterization of a single-chain catalytic antibody
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 4001-4004 (1991)
 MEDLINE 91219500
 FEATURES
 source 1..743
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 gene 1..735
 /gene="7A4-1/212"
 CDS <1..735
 /gene="7A4-1/212"
 /codon_start=1
 /transl_table=1
 /product="7A4-1/212 SCA fusion protein"
 /protein_id="AAA73053.1"
 /db_xref="PID:g207840"
 /db_xref="GI:207840"
 /translation="DVMTPQPSLAMSQKVTMSCKSSQSLNISNKNYLAWVQO
 KQSPKLLIYFASTRSGVDRFTGSGDFTLTISVQAEADQADYFCQOHYRAPR
 TFGGCTKLEIKGTSKSGSEKSGVGLVAPQSLSITCTVSGTSLGRYN
 VHWVQSGPKGLEWGLMIVWGGSTIEYNPALKSRSLSKDNKSKQIFLKNLSLTQDDSA
 MYCYSGYGGDRFYSYWGQGLTVTS"
 misc_feature 1..339
 /gene="7A4-1/212"
 /note="light-chain variable region of mouse antibody
 NPN43C9"
 misc_feature 340..381
 /gene="7A4-1/212"
 /note="synthetic linker"
 misc_feature 382..735
 /gene="7A4-1/212"
 /note="heavy-chain variable region of mouse antibody
 NPN43C9"
 BASE COUNT 195 a 176 c 190 g 182 t
 ORIGIN

Query Match 58.4%; Score 417.2; DB 12; Length 840;
 Best Local Similarity 76.3%; Pred. No. 1.9e-104;
 Matches 563; Conservative 0; Mismatches 148; Indels 27; Gaps 3;

12


```

FT      mat_peptide      1..738
FT      /product='inhibitor of binding og IL-6 to its
FT      receptor'.
FEATURES
Source
1..738
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
187 a 189 c 180 g 182 t
BASE COUNT
ORIGIN

```

| | | | | | | | |
|-----------------------|--------|--------------|--------------|------------|------|--------|------|
| Query Match | 55.5%; | Score | 396.2; | DB | 5; | Length | 738; |
| Best Local Similarity | 75.5%; | Pred. | NO. 1.1e-98; | | | | |
| Matches | 555; | Conservative | 0; | Mismatches | 153; | Indels | 27; |
| | | | | | | Gaps | |

| | | | |
|----|-----|---|-----|
| Qy | 4 | atigtgatgaccagactccaaattctctgtgtatcagcaggagacaggggttaccata | 63 |
| | | | |
| Db | 7 | ATCCTGCTGACACAGTCTCCAAATTCCTGCTGTATCAGCAGGACACAGGTTACCATA | 66 |
| | | | |
| Qy | 64 | acctgcaagggccagtgcagagtgcagtaatgatgtggctgtgtaccacagaagccagg | 123 |
| | | | |
| Db | 67 | ACCTGAAGGCCAGTCAGAGTGTAGTACTGATGAAGTTGTACCAACAGAAGCCAGGG | 126 |
| | | | |
| Qy | 124 | cagtctccgaacactgcgtatatactctgcatactccaatcgtcacactgcagtcgatcgc | 183 |
| | | | |
| Db | 127 | CAGTCTCTAAACTACTGATATATACTATGTCATGCCAATCGCTACACTGGAGTCCCTGATCGC | 186 |
| | | | |
| Qy | 184 | ttcacitggcagtgatgatggacgaatttcactttccaccatcagcactgtgcagctgaa | 243 |
| | | | |
| Db | 187 | TTCACTGGCAGTCGATATGGGACGATTTCACTTTCCACATCAGCATGTGTCAGGCTGAA | 246 |
| | | | |
| Qy | 244 | gacctggcagtttattctgtcagcagattatag-----ctgcgtcggagggggg | 294 |
| | | | |
| Db | 247 | GACCTGGCAGTTTATTTCCTGTGTCAGCAGGATTATAGGTTCTCCATTCCAGTTCGGGTCGGG | 306 |
| | | | |
| Qy | 295 | accaagctggaaataaaagtgaggcgggttcaggcggaggtggcctctgcggtggcgga | 354 |
| | | | |
| Db | 307 | ACAAAGTTGGAAATAAAA---GTCGACAAATCTCTCAGGATCTGGCTCCGAATCCAAAAGC | 363 |
| | | | |
| Qy | 355 | tgcgaggtgcaggtgaaggagtcaggacctgcctggtggcgccctcacagagcctgtcc | 414 |
| | | | |
| Db | 364 | ACCGAGTCAAACTCGAGGAGTCTGGCCCTGGGATTTTGCAGCCCTCCAGACCCCTCAGT | 423 |
| | | | |
| Qy | 415 | atcacttgcactgtctctgggttttcattaaccaattat-----gggtacactgggtt | 468 |
| | | | |
| Db | 424 | CTGACTTGTCTTCTCTGGGTTTCACTGAGCACTTCTGTTATGGTGTGAGCTGGATT | 483 |
| | | | |
| Qy | 469 | cgccagcctccaggaaagggtctgagtggtctgggagtaatatgggctggtggaagcaca | 528 |
| | | | |
| Db | 484 | CGTCAGCCTTCAGGGAAGGGCTCGAGTGGCTGGGCACACATTTATGGGATGATGACAAA | 543 |
| | | | |
| Qy | 529 | aattataattcggctcttatgtccagactgagcatcagcaaggaacaactccaagagccaa | 588 |
| | | | |
| Db | 544 | CACATATAACCCATCCCTGAAGCGCGGCTCACATCTCCAGGATACCTCCACCAACCAG | 603 |
| | | | |
| Qy | 589 | gtttcttataaatgaaactgtgcacaaactgatgcacagccatgtactactgtgcagtt | 648 |
| | | | |
| Db | 604 | GTATTCTCTCAAGATCACCAAGTGTGGACACTGCAGATAGTGCACATATCTTGTGTGTCGA | 663 |
| | | | |
| Qy | 649 | cg-----ggggggtaactacggcctatcctttggactactgggtcgaagaacctca | 699 |
| | | | |
| Db | 664 | AGAACTCTATAGGTAAATTTGGGGGGACTATGCTATGGACTACTGGGGTCAAGGAACCTCA | 723 |
| | | | |
| Qy | 700 | gtcacggctctcctca | 714 |
| | | | |
| Db | 724 | GTACCGTCTCTCTCA | 738 |

| | | | |
|------------|------------------------------------|--------|-----------------|
| RESULT | 7 | | |
| I45910 | | | |
| LOCUS | I45910 | 738 bp | DNA |
| DEFINITION | Sequence 6 from patent US 5639455. | | |
| ACCESSION | I45910 | | |
| | | | PAT 07-OCT-1997 |

| | |
|------------|--|
| NID | G2469875 |
| VERSION | I45910.1 GI:2469875 |
| KEYWORDS | |
| SOURCE | Unknown. |
| ORGANISM | Unknown. |
| REFERENCE | Unclassified. |
| AUTHORS | 1 (bases 1 to 738) |
| TITLE | Shimamura,T., Nakazawa,H. and Hamuro,J. Immunosuppressant |
| JOURNAL | Patent: US 5639455-A 6 17-JUN-1997; |
| FEATURES | : Location/Qualifiers 1..738 /organism="unknown" |
| source | |
| BASE COUNT | 187 a 189 c 180 g |
| ORIGIN | 182 t |

| Query Match | 55.5% | Score 396.2 | DB 5 | Length 738 |
|-----------------------|----------------|--|-----------|------------|
| Best Local Similarity | 75.5% | Pred. No. 1.1e-98 | | |
| Matches 555 | Conservative 0 | Mismatches 153 | Indels 27 | Gaps |
| Qy | 4 | attctgatgccagcagactccaaattcctctgtatcacagagaacaggggtaccata | 63 | |
| Db | 7 | ATCTGTGACACAGTCTCCAAATTCCTGTTGATACAGAGACAGGGTTACCATA | 66 | |
| Qy | 64 | acctgcaagccagctcagagtgtagtaatagtggtctggtaccacaagaagccaggg | 123 | |
| Db | 67 | ACCTGCAAGCCAGTCAGAGTGTGAGTACTGATGTAAGTTGGTACCAACAGAGCCAGGG | 126 | |
| Qy | 124 | cagttccgaaactgctgtatataactctgtacatccaatcgctacactggagttccctgatcgc | 183 | |
| Db | 127 | CAGTCTCCTAACTACTGATATACTATGTCATCCAAATCGCTACACTGGAGTCCCTGTATCGC | 186 | |
| Qy | 184 | ttcaactggcagtggaatggagcggatttcacittcaccatcagcactgtgcagcgtgaa | 243 | |
| Db | 187 | TTCACTGGCAGTGGATATGGACGGATTTCACITTCACCATCAGCACTGTGCAGGCTGAA | 246 | |
| Qy | 244 | gacctggcagtttatctctgcagcaggattatag-----ctgcctcggagggggg | 294 | |
| Db | 247 | GACCTGGCAGTTTATTCTGTGCAGCAGGATTATAGGTCTCCATTACGTTCTGGGCTCGGG | 308 | |
| Qy | 295 | acaaagctggaaataaagggtggagcgggttcaggcggagaggtggctctgcggtggcgga | 354 | |
| Db | 307 | ACAAAGTTGGAAATAAAA---GTGCGACAAATCTCTCAGGATCTGGCTCCGAATCCAAAGC | 365 | |
| Qy | 355 | tcgcaggtgcaggtggaagagtcagagacctggcctgggtgcgccctcacagagcctgcc | 414 | |
| Db | 364 | ACGCAGGTCAAACTCGAGGAGCTCTGGCCCTGGGNATTTGCAGGCCCTCCAGACCCCTCAGT | 423 | |
| Qy | 415 | atcaattgcactgctctgggttttcaataccaattat-----gggtgcactactgggtt | 468 | |
| Db | 424 | CTGACTTGTGTTCTTCTCTGGGTTTTCCTACTGAGCACTTCTGTGATGGGTGTGAGCTGGATT | 483 | |
| Qy | 469 | cgccagcctccaggaaggtctgagctgagtgctgggagtaatatggctggtggaagcaca | 528 | |
| Db | 484 | CGTCAGCCTTCAGGGAAGGCTGTGGAGTGGCTGGCACACATTTATTGGGATGATGACAAA | 543 | |
| Qy | 529 | aattataattcggctcttatgtccagactgagcatcagcaggaactccaagagccaa | 588 | |
| Db | 544 | CACATTAACCCATCCCTGAAGAGCGCGCTCACAATCTCCAAGGATACCTCCACCAACCG | 603 | |
| Qy | 589 | gttttctaaaaatgaacagctctgcaaacctgatgacacagccatgtactactgtgcaggt | 648 | |
| Db | 604 | GTATTCTCAAGATTCACCACTGTGGACACTGTCAGATACTGCCACATACTTCTGTGCTCGA | 663 | |
| Qy | 649 | cg-----ggggggtaactacgcgtatgctttggactactcggggtcgaagaacctca | 699 | |
| Db | 664 | AGAAGTCTCTATGGTATTGGGGGACTATGCTATGCACTACTTGGGGTCAAGGNACCTCA | 723 | |
| Qy | 700 | gtcaccgtctctctca 714 | | |
| Db | 724 | GTCACCGTCTCTCTCA 738 | | |

| QY | 64 | acgtgcaagccagtcagagtgtagtaataatgtagtggcttggtggttaccacagacgagcagg | 123 |
|-------------|-----------|---|-------------|
| Db | 135 | ACTTGAAGCGGAGTCAGGACATTAATAGCTATTTAAAGCTGGATCCAGCAGAACACAGG | 194 |
| QY | 124 | cagtctcgaaactgctgatatctctgcatccaatcgtacactggagtcctgatgcgc | 183 |
| Db | 195 | AAATCTCTAAGACCTCTGATCTACCGTGGAAACAGATTGGTTCCTGGGTCCTCAACAGG | 254 |
| QY | 184 | ttcactggcagtgatggcagcgatttcacattccacatcagcactgtagcagctaaa | 243 |
| Db | 255 | TTGAGTGGCAGTGGATCTGGGCAAGATATTTCTCTCACCATCAGCAGCCTGGAGTATGAA | 314 |
| QY | 244 | gacctggcagtttattctgtcagcaggattatagctcgtcgtggagggggagcccaagctg | 303 |
| Db | 315 | GATGTGGGAGTTTATTATTGTTCTACGGTATGATGAGTTTCCATTACAGTTCGGCTCGGGG | 374 |
| QY | 304 | gaataaaagtgaggcgggttcaggcggagggtggtcgtcgtggcgggtggtcggatcgagctg | 363 |
| Db | 375 | ACAAAGTTGGAAATCAAGCGTGTCTGTCAGAGGCTCTCTCG-----TTCAAGTT | 422 |
| QY | 364 | caggtaagagtcagcagcactggcgtggtggcgcctcagagcgtctccatcactgc | 423 |
| Db | 423 | CAGCTGAGCAGTCAGGACCTGGCCTAGTGCAGCCTTCACAGAGCCTGTCATCACCCTGC | 482 |
| QY | 424 | actgtctcgtgggttttcttaataacaaattatggtgtacactgggttcgcagcctccagga | 483 |
| Db | 483 | ACAGTCTCTGGTTTCTCATTAACCTACCTATGTTGTTACACTGATGTTGCGCAGTCTCCAGGA | 542 |
| QY | 484 | aagggctggagtggtgggagtaataatgggtggtgggaagcacaaaattataatcggct | 543 |
| Db | 543 | AAGGGTCTGGAGTGGGTGGAGTGATATGAGTGGTGGAGCAGACATATAATCAAGT | 602 |
| QY | 544 | cttatgcccagactgagcatcagcagcagcaactcacaagcagcagctttcttaaaatg | 603 |
| Db | 603 | TTATATCCAGCTGAGCATCACAGGACATTCACAGGACCAAGTTTCTTTAAATG | 662 |
| QY | 604 | acagctctgcaaaactgatcacacagcctactgtactactgtgcagctcggggggtaactac | 663 |
| Db | 663 | AACAGTCTGCAAGCTAATGACACAGCCATATATTACTGTGCCA-----GAAATATTGG | 716 |
| QY | 664 | ggtatgcttggtgactactggtgggtcaaggaactcagtcacgctctctca | 714 |
| Db | 717 | GGAACCTCTATGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCA | 767 |
| RESULT | 10 | | |
| LOCUS | SYNMA15C5 | 824 bp ss-DNA | 09-FEB-1994 |
| DEFINITION | | Recombinant molecule with murine monoclonal antibody MA-15C5 variable region, complete cds. | |
| ACCESSION | | M69286 | |
| NID | | 9208807 | |
| VERSION | | M69286.1 | GI:208807 |
| KEYWORDS | | V-region; monoclonal antibody MA15C5. | |
| SOURCE | | Artificial gene DNA. | |
| ORGANISM | | synthetic construct | |
| REFERENCE | | 1 (bases 1 to 824) | |
| AUTHORS | | Holvoet,P., Collen,D., Matthysens,G., Demarsin,E., Gansemans,Y., Stassen,J.-M., Demaeijer,M. and Larice,J. | |
| TITLE | | Characterization of a recombinant single-chain molecule comprising the variable domains of a monoclonal antibody specific for human fibrin fragment D-dimer | |
| JOURNAL | | J. Biol. Chem. 266, 16343-16349 (1991) | |
| MEDLINE | | 91358410 | |
| FEATURES | | | |
| source | | Location/Qualifiers | |
| gene | | 1..824 | |
| sig_peptide | | /organism="synthetic construct" | |
| CDS | | /db_xref="taxon:32630" | |
| | | 12..71 | |
| | | 12..770 | |
| | | /codon_start=1 | |

NID
VERSION M36228.1 GI:194839
KEYWORDS V-region; immunoglobulin heavy chain; processed gene.
SOURCE Mouse (strain BALB/c), cDNA to mRNA, from hybridoma H61-15.

ORGANISM
Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Kavalier, J., Caton, A.J., Staudt, L.M., Schwartz, D. and Gerhard, W.
TITLE A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin

JOURNAL J. Immunol. 145, 2312-2321 (1990)
MEDLINE 90375932
COMMENT Draft entry and computer-readable sequence for [J. Immunol. (1990). In press] kindly submitted by J.Kavaler, 06-JUL-1990.

FEATURES
source location/Qualifiers
1..354 /organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="H61-15"
/tissue_type="hybridoma"
/map="chromosome 12"
1..354 /gene="IgH"
CDS /gene="Igh"
/codon_start=1
/product="immunoglobulin heavy chain V-region"
/protein_id="AAA38067.1"
/db_xref="PIR:I94840"
/translation="QVQLKESGPELVAPQSITCTVSGFSLTSGVHWVRPPKG LEWLGVIWAGGSNTYNSALMSRLSKDINSKVFLKMNSLQDDTTAMYYCARGHYGY YVAMDYWGQGTSVTVS"

BASE COUNT 90 a 89 c 93 g 82 t
ORIGIN

Query Match 45.3%; Score 323.6; DB 12; Length 354;
Best Local Similarity 94.6%; Pred. No. 9.8e-79;
Matches 335; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 358 cagggtcagggtgaagagtcaggacctggctggtgcgcctcacagagcgctgccatc 417
DB 1 CAGGTGCAGTCGAAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTGCCATC 60

QY 418 acttgcaactgtctctgggtttttcattacaaatattggtgtacactgggttcgaccact 477
DB 61 ACTTGCACTGTCTCTGGGTTTTTCATTACCAGCTATGTTGTACACTGGGTTTCGCCACGCT 120

QY 478 ccaggaaaagggtctggagtggctgggagtaatatgggtggtggaagcacacaattataat 537
DB 121 CCAGGAAAGAGTCTGGAGTGGCTGGAGTAATATATGGGCTGGTGAAGCACAAATAAT 180

QY 538 tcgggtcttatgtccagactgagcatcagcagagaactccaagagaccaagtttcttta 597
DB 181 TCGGCTCTCATGTCCAGACTGAGCATCAGAAAGACAACACTCCAAGAGCAAGTTTCTTA 240

QY 598 aaaatgaacagtcctgcaaacactgatgacacagccatgtactactgtgcagtcgggggggt 657
DB 241 AAAATGAACAGTCTCAAACACTGATGACACAGCCATGTACTACTGTGCAGAGGCATTAC 300

QY 658 aactcagcgtctgttttgactactgggttcaaggaacctcagtcacgcgtctcc 711
DB 301 TAGCGGTACTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACGCTGCC 354

RESULT 12
MUSIGHAEI
LOCUS Mouse Ig heavy-chain mRNA V region, partial cds from hybridoma H61-15VH.
DEFINITION
ACCESSION M36228

| | |
|------------|----|
| RESULT | 11 |
| MUSIGHAEI | |
| LOCUS | |
| DEFINITION | |
| ACCESSION | |

LOCUS MUSIGM0520 345 bp mRNA ROD 29-OCT-1994
DEFINITION Mouse IgM chain mRNA VJ4 region, family Q52N.
ACCESSION M76414
NID 9197943
VERSION M76414.1 GI:197943
KEYWORDS J-region: V-region: immunoglobulin heavy chain;
immunoglobulin-kappa.
SOURCE Mus musculus (strain BALB/c, sub_species domesticus) female ascites
CDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 345)
AUTHORS Deng, Y.J., Chua, M.M., Andrews, G.C. and Karush, F.
TITLE Primary B-cell response to neuropeptide Y and bovine pancreatic
polypeptide
JOURNAL Mol. Immunol. 29 (7-8), 847-856 (1992)
MEDLINE 92342150
FEATURES
source
Location/Qualifiers
1..345
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="DSP2.2"
/tissue_type="hybridoma"
/sex="female"
/tissue_type="ascites"
/map="chromosome 12"
1..345
/gene="IgM"
1..345
/partial
/gene="IgM"
/codon_start=1
/product="mu-immunoglobulin"
/protein_id="AA51159.1"
/db_xref="PID:g197944"
/db_xref="GI:197944"
/translation="LESGPGLVAPQSLSITCTVSGFLTSYGVHVRQPPGKGLWL
GVWAGGTNTVNSALMSRLSISKNSKQVFLKNSLQTDTTAMYYCASPILITYYAM
DYWGQTSVTYSS"
4..278

gene
CDS
V_region
N_region
D_segment
N_region
J_segment
BASE COUNT 86 a 90 c 84 g 85 t
ORIGIN
Query Match 44.1%; Score 314.8; DB 12; Length 345;
Best Local Similarity 95.0%; Pred. No. 2.6e-76;
Matches 325; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 373 gagtcaggacctggctgtggtggccctcacagagcctgccatcactgtcactgtctct 432
Db 4 GAGTCAGGACCTGGCCCTGGTGGCGCCCTCACAGAGCCTGCCATCACTTGCACCTGCTCT 63
QY 433 ggggttttcaataacaaataggtgtaactgggttcgcagcctccagaaagggtctg 492
Db 64 GGGTTTTCATTAAACAGCATGATGGTGATACACTGGGTTCGCCAGCCTCCAGAAAGGCTCTG 123
QY 493 gagtgctgggagtaataatgggtgggtggaagcacaaataataatcggtctcttatgtcc 552
Db 124 GAGTGCTGGGAGTATATGGCTGGTGGAGCAGACAAATATATAATTCGCTCTCATGTC 183

QY 553 agactgagcatcagcaaggacaaactccaagagccaaagtttttcttaaaataaacagttgtg 612
Db 184 AGACTGAGCATCAGCAAGACAACTCCAGAGCCAAAGTTTCTTTAAAAATGAACAGTCTG 243
QY 613 caaactgatgacacagccatctactactgtccactcggtgggggtaactacagctatct 672
Db 244 CAAACTGATGACACAGCCATGACTACTGTGCCACCCCTTGATTACTATATGCT 303
QY 673 ttggactactgggtggaaggacccctcagtcaccgtctctctca 714
Db 304 ATGGACTACTGGGTCAGGAACCTCAGTCACCGCTCTCTCA 345
RESULT 13
MUSIGHADX
LOCUS Mouse Ig heavy-chain mRNA V region, partial cds. H220-22VH.
DEFINITION M36217
ACCESSION M36217
NID 9194815
VERSION M36217.1 GI:194815
KEYWORDS V-region; immunoglobulin heavy chain; processed gene.
SOURCE Mouse (strain BALB/c), CDNA to mRNA, from hybridoma H220-22.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kavalier, J., Caton, A.J., Staudt, L.M., Schwartz, D. and Gerhard, W.
TITLE A set of closely related antibodies dominates the primary antibody
response to the antigenic site CB of the A/PR/8/34 influenza virus
hemagglutinin
JOURNAL J. Immunol. 145, 2312-2321 (1990)
MEDLINE 90375932
COMMENT Draft entry and computer-readable sequence for [J. Immunol. (1990)
In press] kindly submitted
by J.Kavalier, 06-JUL-1990.
FEATURES
source
Location/Qualifiers
1..360
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="H220-22"
/tissue_type="hybridoma"
/map="chromosome 12"
1..360
/gene="IGH"
1..360
/gene="IGH"
/codon_start=1
/product="immunoglobulin heavy chain V-region"
/protein_id="AAA38056.1"
/db_xref="PID:g194816"
/db_xref="GI:194816"
/translation="QVQLKESGPGCLVAPQSLSITCTVSGFLTSYGVHVRQPPGKGL
LEWLGVIWAGGTNTVNSALMSRLSISKNSKQVFLKNSLQTDTTAMYYCARGHYG
DGYVMDYWGQTSVTYS"
BASE COUNT 91 a 89 c 96 g 84 t
ORIGIN
Query Match 43.5%; Score 310.8; DB 12; Length 360;
Best Local Similarity 93.6%; Pred. No. 3.2e-75;
Matches 337; Conservative 0; Mismatches 17; Indels 6; Gaps 1;
QY 358 caggtgcagggtgaaggagtcaggacctggctgtggccctcacagagctgtccatc 417
Db 1 CAGGTGCAGGCTGAAGAGTCAGGACCTGGCTGGTGGCCCTCACAGAGCTGTGTCATC 60
QY 418 acttgacactgtctctgggttttcttaacaaattatggtgtacactgggttcgcagcct 477
Db 61 ACTTGCACCTGTCTCTGGGTTTTCATTAACAGCATGATGTTGATACACTGGTTCGCCAGCT 120
QY 478 ccaggaaaagggtctggagtggtgggtgagtaataatgggtggtggaagcacaaataataat 537

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 30, 1999, 12:44:32 : Search time 383.27 Seconds
(without alignments)
4382.041 Million cell updates/sec

Title: US-08-940-544-3
Perfect score: 717
Sequence: 1 caggtgaactgcagcagtc.....cacagttggaataaaacgg 717

Scoring table: IDENTITY_NUC

Searched: 3032314 seqs, 1171202697 residues

Database :

EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 277 | 38.6 | 396 | 37 | AA691311 vs14f01.r |
| 2 | 203.8 | 28.4 | 344 | 35 | AA592800 vo25g11.r |
| 3 | 197 | 27.5 | 380 | 49 | AI641928 vq50c10.y |
| 4 | 182.4 | 25.4 | 293 | 51 | AV052316 AV052316 |
| 5 | 182.2 | 25.4 | 310 | 43 | AI180569 uc70f09.r |
| 6 | 180.2 | 25.1 | 461 | 43 | AI201426 q573c06.x |
| 7 | 178.6 | 24.9 | 453 | 62 | AI923980 wh49d07.x |
| 8 | 177.2 | 24.7 | 528 | 60 | AI799757 wc37g12.x |
| 9 | 173.6 | 24.2 | 294 | 41 | AI007196 ua73g04.r |
| 10 | 171.2 | 23.9 | 297 | 51 | AV052647 AV052647 |
| 11 | 170.6 | 23.8 | 488 | 60 | AI791363 ob68a09.y |
| 12 | 159.8 | 22.3 | 387 | 35 | AA569186 nm30d10.s |
| 13 | 126.6 | 17.7 | 393 | 20 | T27593 EST100653 H |
| 14 | 122.6 | 17.1 | 379 | 31 | AA291381 zt44g02.r |
| 15 | 122.2 | 17.0 | 345 | 37 | AA710291 vt53a04.r |
| 16 | 121.8 | 17.0 | 400 | 34 | AA456778 2w27f11.r |
| 17 | 119.4 | 16.7 | 398 | 33 | AA423447 ve80a03.r |
| 18 | 119.2 | 16.6 | 336 | 40 | AA987559 or83g01.s |
| 19 | 119.2 | 16.6 | 470 | 44 | AI306478 qnl3b10.x |
| 20 | 119.2 | 16.6 | 640 | 50 | AI697483 we14a04.x |
| 21 | 119 | 16.6 | 413 | 31 | AA301347 EST14279 |
| 22 | 118.4 | 16.5 | 370 | 31 | AA295093 EST100400 |
| 23 | 117.2 | 16.3 | 431 | 37 | AA710970 vt93b09.r |
| 24 | 115.6 | 16.1 | 328 | 31 | AA295311 EST10047 |
| 25 | 115.4 | 16.1 | 345 | 32 | AA335086 EST39457 |
| 26 | 114.6 | 16.0 | 402 | 48 | AI549800 ve80a03.y |
| 27 | 114.4 | 16.0 | 279 | 31 | AA300571 EST13661 |
| 28 | 112.8 | 15.7 | 349 | 32 | AA379044 EST91999 |
| 29 | 112.6 | 15.7 | 471 | 43 | AI233978 EST230666 |
| 30 | 110 | 15.3 | 382 | 31 | AA295786 EST100987 |
| 31 | 110 | 15.3 | 333 | 31 | AA300732 EST13847 |
| 32 | 109.8 | 15.3 | 463 | 22 | R69532 yj82d09.r1 |
| 33 | 109.6 | 15.3 | 420 | 42 | AI120005 uc25e04.r |
| 34 | 109.4 | 15.3 | 253 | 31 | AA295941 EST101165 |
| 35 | 109 | 15.2 | 395 | 20 | T29112 EST69384 Hu |
| 36 | 109 | 15.2 | 335 | 31 | AA318377 EST20820 |
| 37 | 108.2 | 15.1 | 352 | 31 | AA301261 EST14181 |
| 38 | 107.8 | 15.0 | 363 | 32 | AA367405 EST78511 |
| 39 | 106.8 | 14.9 | 403 | 32 | AA345486 EST51505 |
| 40 | 106.6 | 14.9 | 291 | 32 | AA377074 EST89603 |
| 41 | 105.8 | 14.8 | 209 | 20 | T28938 EST61186 Hu |
| 42 | 104.4 | 14.6 | 287 | 20 | T27609 EST101034 H |
| 43 | 104.4 | 14.6 | 381 | 20 | T29114 EST69430 Hu |
| 44 | 104.4 | 14.6 | 823 | 29 | AA170256 ms87g10.r |
| 45 | 104.2 | 14.5 | 383 | 20 | T29916 EST99871 Hu |

ALIGNMENTS

```

RESULT 1
AA691311 396 bp mRNA EST 16-DEC-1997
LOCUS vs14f01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
DEFINITION clone IMAGE:1138201 5', similar to gb:X67211 M.musculus rearranged
immunoglobulin kappa light chain (MOUSE);, mRNA sequence.

ACCESSION AA691311
NID g2692247
VERSION AA691311.1 GI:2692247
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 396)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397933.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:619473
Seq primer: -28ml3 rev2 ET from Amersham.
Location/Qualifiers
1. 396
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/map="11 q13.4-q23.1; 972B12"
/clone="IMAGE:1138201"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'GTGTACGAATCTGAAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGATCCTTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library constructed by Bob Barstead."
BASE COUNT 104 a 107 c 91 g 94 t
ORIGIN

Query Match 38.6%; Score 277; DB 37; Length 396;
Best Local Similarity 95.0%; Pred. No. 3.4e-69;
Matches 286; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 398 acatcgactcactcagctccagcaaatcatgtctgtatcccgaggagagctcacc 457
Db 93 AAATTGTTCTCACCAGCTCTCAGCAATCATGTCTGCATCTCCAGGGAGAGGTACCA 152
Oy 458 tgaactgcagtgccagctcaagtataaagttacatgcactggtaccagcagaagcctgtcca 517
Db 153 TGACCTGCAGTCCGAGCTCAAGTGTAACTTACATGCATCTGTACCAGCAGAGTCAGGCA 212

QY 518 cctcccccaaaagatggatttatgacacatccaaactggcttcttgagagtcctcgtcgct 577
Db 213 CTTCCCAAAAGATGGATTATGACACATCCAAACTGGCTTCTGGAGTCCCTGCTGCT 272
OY 578 tcagtgcactgggtctgggacaccttattctcacaatcacacacatgagagctgtag 637
Db 273 TCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGATGAGGCTGAAG 332
OY 638 atgtgcacttattactgccatcagcggagtagttacccgcctcacgttcggtgctggga 697
Db 333 ATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAAACCGCTCACGTTCCGTCCTGGGA 392
OY 698 c 698
Db 393 C 393

RESULT 2
AA592800 344 bp mRNA EST 17-SEP-1997
LOCUS vo25g11.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
DEFINITION clone IMAGE:1050980 5', similar to gb:X67211 M.musculus rearranged
immunoglobulin kappa light chain (MOUSE);, mRNA sequence.

ACCESSION AA592800
NID g2406463
VERSION AA592800.1 GI:2406463
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 344)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Apr 18, 1995 this sequence version replaced gi:775310.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:582556
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 344
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1050980"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'GTGTACGAATCTGAAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGATCCTTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library constructed by Bob Barstead."
FEATURES
source

```

FEATURES
SOURCE

| | |
|-----------|---|
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; |
| REFERENCE | Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | 1 (bases 1 to 461) |
| TITLE | NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap . |
| | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), |
| JOURNAL | Tumor Gene Index |
| COMMENT | Unpublished (1997) |
| | On Jul 25, 1995 this sequence version replaced gi:637932. |

```

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1530
Email: Robert_Strausberg@nih.gov
unknown library type
Insert Length: 608 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .461
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1943722"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonELDS
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
93 a 132 c 106 g 130 t
BASE COUNT
ORIGIN

```

| | | | | | | | |
|-----------------------|---|---|------------|------------|-------------|--------|------|
| Query Match | 25.1% | Score | 180.2; | DB | 43; | Length | 461; |
| Best Local Similarity | 74.9%; | Pred. | No. 2e-41; | | | | |
| Matches | 224; | Conservative | 0; | Mismatches | 75; | Indels | 0; |
| | | | | | | | |
| QY | 1 | caggtgaactcgcgcagctcagacactggaactggtgnagcctgggcttcagtgaaacata | 60 | | | | |
| | | | | | | | |
| Db | 438 | CAGGTGCAGCTGGTGCAGTCTGGGGCTCAGGTGAAGAGGCTGGGGCTCAGTGAAGGTC | 379 | | | | |
| | | | | | | | |
| QY | 61 | tctctcaagacttcttggnacaaatctcactgaatacacaccatgcactgggtgaagcagagc | 120 | | | | |
| | | | | | | | |
| Db | 378 | TCTGTGAAGCGTTCTGGATACATCTTCACCGACTACTATATGCACCTGGGTGCGACAGGCC | 319 | | | | |
| | | | | | | | |
| QY | 121 | catggaaagaccttcagtgagattgaggtattaactcctaacaatggtgtactaacctac | 180 | | | | |
| | | | | | | | |
| Db | 318 | CTTGGACACAGAGCTTGGGTGGATGGGAGGATCAACCTTAACGTGTTGCACAAACTAT | 259 | | | | |
| | | | | | | | |
| QY | 181 | aagcagaagttcaaggcgcaaggccacattgactgtacagaagttctccagcacagcctac | 240 | | | | |
| | | | | | | | |
| Db | 258 | GCACAGAAGTTTCAGGGCAGAGTCACCATGCCAGGGACACGTCCTCATCAGCACAGCCTAC | 199 | | | | |
| | | | | | | | |
| QY | 241 | atgagactccgagcctgaacatctgaggaattctgcagctattactctgaagaagatac | 299 | | | | |
| | | | | | | | |
| Db | 198 | ACGGAGCTGAGCAGCCTGAGATCTGAGGACACGCGCCACGTATTACTGTGCGAGAGACAC | 140 | | | | |
| | | | | | | | |
| RESULT | 7 | | | | | | |
| AI923980/c | | | | | | | |
| LOCUS | AI923980 | 453 bp | mRNA | EST | 02-SEP-1999 | | |
| DEFINITION | wn49607.x1 NCI_CGAP_Ju19 Homo sapiens cDNA clone IMAGE:2448781 3' | | | | | | |
| | similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); | | | | | | |
| | mRNA sequence. | | | | | | |
| ACCESSION | AI923980 | | | | | | |
| NID | q5659944 | | | | | | |

AI923980.1 GI:5659944
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3137502.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

Seq primer: -40UP from Gibco
High quality sequence stop: 44.
Location/Qualifiers
1. .453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="10"
/clone="IMAGE:2448781"
/clone_lib="NCI_CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- ClgI(GT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaudo."
91 a 132 c 97 g 133 t

```

| Query Match | 24.9% | Score 178.6 | DB 62 | Length 453 |
|-----------------------|----------------|--|----------|------------|
| Best Local Similarity | 74.6% | Pred. No. 5.6e-41 | | |
| Matches 223 | Conservative 0 | Mismatches 76 | Indels 0 | Gaps 0 |
| QY | 1 | caggtgaaactgcagcagctcagcagctgaactggtgnagcctgggcttcagtgaagata | 60 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 445 | CAGGTCAGCTGGTACAGTCTGGGGTGAGGTGAAAGAAGCCTGGGGCTCAGTGAAGGTC | 386 | |
| | | | | |
| | | | | |
| | | | | |
| QY | 61 | tctgtcaagactcttctggnacaattcactggaatacaccaatgcactgggtggaagcagagc | 120 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 385 | TCCTGCAAGGTTTCCGGATACACCCCTCACTGAATATCCATGCACCTGGGTGCGACAGGCT | 326 | |
| | | | | |
| | | | | |
| | | | | |
| QY | 121 | catggaagagccttcagtggtattgaggtattaaacctcaacaattggtggtactaactac | 180 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 325 | CTGTGAAAAGGGCTTGAGTGGATCGGGAGGTTTTGTGCTTGAAATGCTGAAAACATCTAC | 256 | |
| | | | | |
| | | | | |
| | | | | |
| QY | 181 | aagcagaagttcaagggcgaagccacattgactgtgagacaagttcctccagcagcagctac | 240 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 265 | GCACAGAAGTTTCCAGGGCAGAGTCACCATCGACGAGGACACATCTACAGACACAGGCTAC | 206 | |
| | | | | |
| | | | | |
| | | | | |
| QY | 241 | atgaggtctccagcagcctgcacatctgaggtattcgagctctattactctgtgcaagagatac | 299 | |


```
BASE COUNT      85 a      68 c      77 g      64 t
ORIGIN

Query Match      24.2%; Score 173.6; DB 41; Length 294;
Best Local Similarity 78.6%; Pred. No. 1.2e-39;
Matches 231; Conservative 0; Mismatches 61; Indels 2; Gaps 2;

QY 3 ggtgaagactcagcagtcaggacac-tggtgnagcctgggcttcagtgaaagatat 61
Db 1 GGTCCAACTCAGCAGCCTGGGCTGAGCTGGTGAGGACTGGGTCTTCAGTGAAGCTG 60

QY 62 cctgcaagactctgganacaaattcaactgaatacacaccatgcactgggtgaagcagagcc 121
Db 61 ACTGCAAGGACTCAGGCTACACCTTCACCAGCTACTGGATGCATTGGTGAAGCAGAGGC 120

QY 122 atggaagagccttgatggatggaggtatttaactccatacaatgggtgactaactaca 181
Db 121 ATATACAGGCATTGATGATGGTGAACATTGACCTTCAGCTATAGTGGAACTCACTACA 180

QY 182 agcagaagtccaagggcgaagccacattgactgtagacaaagtccctccagcagcctaca 241
Db 181 ATCAAAAGTTCAAGGACAAGGCCACATTTGACTGTAGACAAATTCACCAGCACAGCCTACA 240

QY 242 -tgagctccgagcctgacatctgagattctgcagctctattactctgcaaga 294
Db 241 GTGAGCTCAGCAGACTGACATCTGAGGACTCTGCGGTCTATTACTGTGTGCAAGA 294

RESULT 10
AV052647      297 bp      mRNA      EST      22-JUN-1999
LOCUS      AV052647 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION      clone 1810022L19, mRNA sequence.
ACCESSION      AV052647
NID      95138419
VERSION      AV052647.1 GI:5138419
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 297)
Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
Sugahara,Y., Suzuki,H., Suzuki,H., Tateo,M., Tomaru,Y.,
Tomimaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
RIKEN Mouse ESTs
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced gi:3246882.

Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@r.riken.go.jp
Thermolabile and thermostable activation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..297
/organism="Mus musculus"
/strain="C57BL/6J"

/db_xref="taxon:10090"
/clone="1810022L19"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"

BASE COUNT      79 a      54 c      75 g      89 t
ORIGIN

Query Match      23.9%; Score 171.2; DB 51; Length 297;
Best Local Similarity 75.2%; Pred. No. 5.9e-39;
Matches 212; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 9 actgcagcagtcaggacacctgcactgaatacacaccatgcactgggtgaagcagagccatggaaa 128
Db 16 AATGCAGCAGTTGGAGCTGAGTGGTGACACCCCGGGCTTTAGTGAAGCTGCTTGCAA 75

QY 69 gactctggaanacaaattcaactgaatacacaccatgcactgggtgaagcagagccatggaaa 128
Db 76 GGTTCCTGGCTACACTTTCACCTGAATATATATACACTGGGTAAAGCAGAGTTTGGACA 135

QY 129 gagccttggtgagtgagggtatttaactccatacaatggtggtactaactacaagcagaaa 188
Db 135 GGGCCTTGAGTGGATGGCTGCTGTTTACCTGGGAAGTGGTAGTATAAATTACAAATGAGAA 195

QY 189 gttcaagggcgaagggcaccattgactgtagacaaagtcctccagcacagcctacatgagct 248
Db 196 TTTCAACGACGAAGGCCACATTCGACTCGGACAAATCTTCCAGCACAGTTTATATGAACT 255

QY 249 ccgcagcgtgacatctgaggattctgcagctctattactctgtgc 290
Db 256 TAGTAGATTGACATCAGAGAAATTTGCGGTCTATTTTGTGC 297

RESULT 11
AV052647      488 bp      mRNA      EST      02-JUL-1999
LOCUS      AV052647 oh68a09.y5 NCI-CGAP_K1d5 Homo sapiens cDNA clone IMAGE:1472152 5'
DEFINITION      similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN);,
mRNA sequence.
ACCESSION      AV052647
NID      95339079
VERSION      AV052647.1 GI:5339079
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 488)
NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3138279.
Other ESTs: oh68a09.x5
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
```


GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 12:44:34 ; Search time 39.3 Seconds
(without alignments)
1929.363 Million cell updates/sec

Title: US-08-940-544-3
Perfect score: 717
Sequence: 1 caggtagaactgcagcagtc.....cacagttggaataaaacgg 717

Scoring table: IDENTITY_NUC
Searched: 195662 seqs, 52875861 residues
D: se : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS9_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 521.6 | 72.7 | 726 | 3 | US-08-553-497A-25 |
| 2 | 517.4 | 72.2 | 1797 | 2 | US-08-463-163-2 |
| 3 | 512.6 | 71.5 | 726 | 3 | US-08-553-497A-27 |
| 4 | 507.2 | 70.7 | 810 | 4 | US-08-652-507-1 |
| 5 | 504.6 | 70.4 | 732 | 3 | US-08-553-497A-19 |
| 6 | 503 | 70.2 | 732 | 3 | US-08-553-497A-21 |
| 7 | 501.6 | 70.0 | 1679 | 3 | US-08-661-052-15 |
| 8 | 487 | 57.9 | 738 | 3 | US-08-553-497A-23 |
| 9 | 409.2 | 57.1 | 738 | 4 | US-08-956-047-24 |
| 10 | 397 | 55.4 | 711 | 3 | US-08-190-199A-64 |
| 11 | 386.8 | 53.9 | 708 | 3 | US-08-190-199A-60 |
| 12 | 381 | 53.1 | 720 | 4 | US-08-800-198-7 |
| 13 | 368.2 | 51.4 | 831 | 3 | US-08-403-853-17 |
| 14 | 360.8 | 50.3 | 756 | 4 | US-08-797-689-17 |
| 15 | 348.8 | 48.6 | 2012 | 1 | US-08-235-838-15 |
| 16 | 348.8 | 48.6 | 2012 | 4 | US-08-465-473B-15 |
| 17 | 347 | 48.4 | 748 | 1 | US-08-235-838-10 |
| 18 | 347 | 48.4 | 748 | 4 | US-08-465-473B-10 |
| 19 | 325.6 | 45.4 | 797 | 3 | US-08-894-922A-13 |
| 20 | 325 | 45.3 | 672 | 3 | US-08-190-199A-62 |
| 21 | 324 | 45.2 | 891 | 3 | US-08-894-922A-9 |
| 22 | 314.8 | 43.9 | 669 | 3 | US-08-190-199A-66 |
| 23 | 314.4 | 43.8 | 858 | 4 | US-08-428-257A-71 |
| 24 | 299 | 41.7 | 2233 | 1 | US-08-235-838-6 |
| 25 | 299 | 41.7 | 2233 | 4 | US-08-235-838-13 |
| 26 | 299 | 41.7 | 2233 | 4 | US-08-465-473B-6 |
| 27 | 299 | 41.7 | 2012 | 1 | US-08-465-473B-13 |
| 28 | 298.8 | 41.7 | 748 | 4 | US-08-235-838-4 |
| 29 | 298.8 | 41.7 | 748 | 4 | US-08-465-473B-4 |
| 30 | 295.8 | 41.3 | 793 | 4 | US-08-956-047-29 |
| 31 | 291.8 | 40.7 | 723 | 1 | US-07-988-430-93 |
| 32 | 291.8 | 40.7 | 723 | 2 | US-08-425-336-90 |
| 33 | 291.8 | 40.7 | 723 | 2 | US-08-488-113B-90 |
| 34 | 291.8 | 40.7 | 723 | 2 | US-08-477-484B-90 |
| 35 | 291.8 | 40.7 | 723 | 3 | US-08-646-360-90 |
| 36 | 291.8 | 40.7 | 723 | 5 | PCT-US92-09487-93 |
| 37 | 280.6 | 39.1 | 354 | 3 | US-08-491-845-5 |

ALIGNMENTS

RESULT 1
US-08-553-497A-25
; Sequence 25, Application US/08553497A
; Patent No. 5844093
; GENERAL INFORMATION:
; APPLICANT: KETLEBOROUGH, C. A.
; APPLICANT: BENDIG, MARY M.
; APPLICANT: ANSELL, KEITH H.
; APPLICANT: GUSOW, DETLEF
; APPLICANT: ADAN, JAUME
; APPLICANT: MITJANS, FRANCES
; APPLICANT: ROSELL, ELISABET
; APPLICANT: BLASCO, FRANCESC
; APPLICANT: PIJLATS, JAUME
; TITLE OF INVENTION: ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,497A
; FILING DATE: 17-NOV-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP95/00978
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94104160.0
; FILING DATE: 17-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94118970.6
; FILING DATE: 03-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HAMLET-KING, DIANA
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1726
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEC ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:

Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 18, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 6, Appli

```

; ORGANISM: mouse
; STRAIN: Balb/c
; TISSUE TYPE: splenocytes
; IMMEDIATE SOURCE:
; CLONE: 5 F 1 (single-chain Fv, heavy, light chain,
; CLONE: linker)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..726
; US-08-553-497A-25

Query Match 72.7%; Score 521.6; DB 3; Length 726;
Best Local Similarity 84.4%; Pred. No. 1.4e-138;
Matches 613; Conservative 0; Mismatches 101; Indels 12; Gaps 2;

QY 1 cagtgaaactgcagcagtcaggaccctgaactgtgagcctggggttcagtgaagata 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 CAGGTGAACATGCAAGGAGTCGTGGGGCTGAACCTGGTGAAGCCTGGGGCTTCAGTGAAGTTG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 tctctgaagactctctgganacaaatctcactgaatcacaccatgcactgggtggaagcagag 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 TCTGTCAAGGCTTCGGCTACACCTTCACCAGCCACTGGATGCATGGTGGTGAACGACGAGG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 catggaagagccttgagtgagtgattgaggtattaatcctaacaattggtggtactaac 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 GCTGCACAAGCCCTTGAGTGGATCGGAGAGATTAAATCCAGAACGGCGCCTACTAACTAC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 181 aagcagaagttcaaggcgcaagccacattgactgtagaagtcctccagcacagcctac 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 AATGAGAAATTCAGAGCAAGGCCACACTGACTGTAGACAAATCCTCAGACACAGCCTAC 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 241 atgagcttcgcagccctgacatctgagattcttcagctcttactgtgcaag-agatac 299
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 ATGCAACTCAGCACCCCTGACATCTGAGACTCTCGGGTCTATTACTGTGCCAGTCGGGAC 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 300 taagggtccgg-----ttgttactgggtccaaaggagcaaggtcacogtctcc 348
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 TATGATTACGACGACGGGTACTTTGACTACTTGGGGCCAAAGGACAAAGGTACCGTCTCTCC 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 349 tcagtgaggcggttcaggcgaggtgctctggcggtggcgatcgacatcgagctc 408
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 TCAGGTGGCGGTGCTCGGGCGGTGGTGGTTCGGGTGGCGCGGATCTGCATTGAGCTC 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 409 actcagctctccagcaatcatgtctgcactccaggggagaaaggtccacatgacctgcagt 468
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 ACCAGTCTCCAACAATCATGTCTGCATCTCCAGGGGAGAAAGTACCATTGACCTGAGT 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 469 ggcagctcaagtaataagttacatgcactgggtaccagcagaagcctgtcacctcccccaaa 528
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 481 GACAGCTCAAGTGTAAAGTTACACGTACTGTGTACCAGCAGAAAGACAGGATCCTCCCCAGA 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 529 agatggatttatgacacatccaaactggctctggagtcctctgcgttcagtgagcagt 588
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 541 CTCTGTATTTATGACACATCCAACTCGCTTCTGGAGTCCCTGTTCGCTTCAGTGGCAGT 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 589 gggctcgggacctcttattctcacaaatcagcagcatgaagcgttagatgctgcacct 648
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 601 GGGTCTGGGACCTCTTACTCTCTCAAAATAGCCGAATGGAGGCTGAAGAATGCTGCCACT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 649 tattactgccatcagcggagtagttacccgtcacgttcggtgctgggacacagtgtgaa 708
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 661 TATTACTGCCAGCAGTGGAGTAGTTACCCGCTCAGTTCGGTTCGGTGTGGGACCAAGCTGGAA 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 709 ataaaa 714
| | | | |
DB 721 ATRAAA 726
| | | | |

RESULT 2
US-08-463-163-2
; Sequence 2, Application US/08463163
; Patent No. 5696237

```


Db 482 ACAGCTCAAGTGAAGTACATGCTACTGCTACAGCAGCAAGACAGGATCTCCCCACAC 541
Qy 530 gatggattatgacacatcccaactgggtctctggagtcctcgtcgtcctcagtgagcagtg 589
Db 542 TTCTGATTATGACACATCCAACTGGCTTCTGGAGTCCCTGTTGCTTCAAGTGGCAGTG 601
Qy 590 ggtctggagactcttattctctcaaatcagcagcatggaggtctgagatcctccactt 649
Db 602 GGTCTGGAGACTCTTACTCTCTCAATCAGCCGAATGAGGCTGAAGATGCTGCCACTT 661
Qy 650 attactgcatcagcggagtgattaccgctcagcttcggtcgtcgtgagacacagttgaaa 709
Db 662 ATTACTGCCAGAGTGGAGTAGTTACCCGCTCACGTTCCGTTGCTGGGACCAAGCTGGAAA 721
Qy 710 taaaa 714
Db 722 TAAAA 726

REPT 4

US-08-652-507-1

; Sequence 1, Application US/08652507

; Patent No. 5876691

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: ANTIBODY AGAINST CARCINOEMBRYONIC ANTIGEN

; TITLE OF INVENTION:

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon & Vanderhye, P.C.

; STREET: 1100 No. 5876691th Glebe Road, 8th Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/652,507

; FILING DATE: 02-Jul-1996

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Arthur R. Crawford

; REGISTRATION NUMBER: 25,327

; REFERENCE/DOCKET NUMBER: 117-211

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-816-4000

; TELEFAX: 703-816-4100

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 810 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-652-507-1

Query Match 70.7%; Score 507.2; DB 4; Length 810;
Best Local Similarity 82.2%; Pred. No. 1.6e-134;
Matches 597; Conservative 0; Mismatches 120; Indels 9; Gaps 1;

Qy 1 caggtgaaactgcagcagtcaggacccgaactggtgagcctgggttcagtgaaagata 60
Db 79 CAGGTGAAACTGCGAGCTGCGGCGAGAACTTGTGAGGTGAGGACCTCAGTCAAGTTG 138
Qy 61 tctgcaagacttctgganacaaattcactgaatacacaccatgctcagctgggtgaagcagc 120
Db 139 TCCTGCACAGCTTCTGCTTCAACATTAAGACTTCCTATATGCACTGGTTGAGGCAGGGG 198

Qy 121 catggaagagccttgagtgattggaggtatttaactcctaacaatggtggtactaactac 180
Db 199 CCTGAACAGAGCCCTGGAGTGGATTGGATGATGATCTCCTGAGAAATGGTGATGAATAT 258
Qy 181 aacacaaagtccaagggcaagccacattgactgtagaagaagtcctccagcagcagctac 240
Db 259 GCCCGAAGTTCCAGGGCAAGGCCACTTTTACTACAGACACATCTCTCCAAACACAGCCTAC 318
Qy 241 atggagctccgcagcctgacatctgaggtattctgcagttctattactgtgcaagagatact 300
Db 319 CTCAGCTCAGCAGCCTGACATCTGAGGACACTGCCGCTCTATTATTGTAATGAGGGGACT 378
Qy 301 -----acggtcccggtttgttactgggtccaaaggaccacggtcaccgtctctctca 351
Db 379 CCGACTGGGCCGTACTACTTTGACTACTGTGGGCGCAAGGGACACCGGTCAACCGTCTCTCA 438
Qy 352 ggtggaggggttcaggcgaggtggctctggcgggtggcgagatcgagatcgagctcaact 411
Db 439 GGTGGAGGGCGGTTACAGCGGAGGTGGCTCTGGCGGTGGCGGATCAGAAAAATGTGCTCAOC 498
Qy 412 cagtcctcagcaaatcatgtctgcatctccaggggagaaagtcaccatgacctgcagtggc 471
Db 499 CAGTCTCCAGCAANTCATGTCTGCATCTCCAGGGGAGAGAGGTCAACATAACTCGAGTGCC 538
Qy 472 agtcaagataaagttacatgcactggtaccagcagagaagcctgtcacctcccccaaaaga 531
Db 559 AGTCAAGTGAAGTTACATGCACTGGTTCCAGCAGAAAGCAGGCACTTCTCCCAAACTC 618
Qy 532 tggattatgacacatccaaactggcttctggaatccctcgtcgtcagtgagcagtgagg 591
Db 619 TGGATTATAGACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGA 678
Qy 592 tctgggacctcttattctcacaatcagcagcatggagctgtagatgctgcccacttat 651
Db 679 TCTGGGACCTCTTACTCTCTCAATCAGCCGAATGGAGGCTGAAGATGCTGCCACTTAT 738
Qy 652 tactgcatcagcggagtagttaccgcgctcaactcgttgcgtgggacacagtgaggaaata 711
Db 739 TACTGCCAGCAAGAGGAGTAGTTACCCACTCACGTTCCGTTGCTGGCAGTGGAGCTG 798
Qy 712 aaacgg 717
Db 799 AAACGG 804

RESULT 5

US-08-553-497A-19

; Sequence 19, Application US/08553497A

; Patent No. 5844093

; GENERAL INFORMATION:

; APPLICANT: KETTLERBOROUGH, C. A.

; APPLICANT: BENDIG, MARY M.

; APPLICANT: ANSELL, KEITH H.

; APPLICANT: GUSSOW, DETLEF

; APPLICANT: ADAN, JAUME

; APPLICANT: MITJANS, FRANCES

; APPLICANT: ROSELL, ELISABET

; APPLICANT: BLASCO, FRANCES

; APPLICANT: PIULATS, JAUME

; TITLE OF INVENTION: ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR

; ANTIBODIES

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.

; STREET: 2200 CLARENDON BLVD. SUITE 1400

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: US

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

Db 362 CAGGTGGCGTGGCTGGCGGTGGTGGTGGCGCGGATCTGGATCTGACATTG 421
Qy 404 agctcactcagctccagcaaatcgtctgcatctccagggagaggtcaccatgacct 463
Db 422 AGCTACCAGCTCCAAATCATCTGCAATCTCCAGGGGAGAGGTCAACATGACCT 481
Qy 464 gcagtgagcgtcaagataaagtacatgcactgggtaccagcagaagcgtgcacctccc 523
Db 482 GCAGTGACAGCTCAAGTGAAGTTACATGACTGTTACCGAGAGAGCCAGGATCTCTCC 541
Qy 524 ccaaaagatgattatgatcacatcaaaactggctctgaggtccctgctcctcaagt 583
Db 542 CCAGACTCTCTGATTTATGACATCAATCAATCAATCAATCAATCAATCAATCAAT 601
Qy 584 gcagtgagcgtcagcgtccttattctcacaatcagcagcagcagcagcagcagcag 643
Db 602 GCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAATCAATCAATCAATCAAT 661
644 ccacttattactgcccacagcagcagcagcagcagcagcagcagcagcagcagcag 697
Db 662 CCAGTATTACTGCCAGCAGTGGAGTAGTTACCCACCATGTACACGTTCCGGAGGGGGA 721
Qy 698 ccagtggaataaaaa 714
Db 722 CCAAGCTGGAATAAAA 738

RESULT 9
US-08-956-047-24
; Sequence 24, Application US/08956047
; Patent No. 5882924
; GENERAL INFORMATION:
; APPLICANT: Fritz, Hans-Joachim
; APPLICANT: Hennecke, Frank
; APPLICANT: Kolmar, Harald
; TITLE OF INVENTION: Genetic Selection, by Means of Signal
; TITLE OF INVENTION: Transduction in Microorganisms, of Proteins Which are
; TITLE OF INVENTION: Capable of Ligand Binding
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESS: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,047
; FILING DATE: 22-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/257,669
; FILING DATE: 08-JUN-1994
; APPLICATION NUMBER: DE P 43 19 296.3
; FILING DATE: 10-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pohlman, Sandra M.
; REGISTRATION NUMBER: P39,691
; REFERENCE/DOCKET NUMBER: 05552.1368-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; OTHER INFORMATION: /note= "Molecule 1-720 encodes a
; OTHER INFORMATION: peptide."
US-08-956-047-24

Query Match 57.1%; Score 409.2; DB 4; Length 738;
Best Local Similarity 74.3%; Pred. No. 7.4e-107;
Matches 532; Conservative 0; Mismatches 175; Indels 9; Gaps 1;

Qy 2 agtggaactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 61
Db 14 ACCTACAGCTGGTGGAGCTTGGGGAGGCTTGTGCGAGCTGGAGGGTCCCGGAACCT 73
Qy 62 cctgcaagactcttgannacaaattcactgaatacacaccatgcactgggtgagcagcag 121
Db 74 CCGTGCAGCTCTGGATTCACTTTTCAGTAGCTTTTGGAAATGCACCTGGCTTCGTACGGCTC 133
Qy 122 atggaagaagccttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 181
Db 134 CAGAGAAGGGGCTGGAGTGGGTGCGATATATAGTAGTGGCAGTAGTACTATCTACTATG 193
Qy 182 agcagaagttcaaggcgaaggccacacattgactgttagacaagtcctccagcagcagc 241
Db 194 CAGACACAGTGAAGGGCGGATTCACCATCTCCAGAGACAATCCCAAGAACACCCCTGTTCC 253
Qy 242 tggagctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 301
Db 254 TGCAGATGACCAAGCTAAAGTCTGAGGACACGGCCATGATTACTGTGCAAGAGATTAGG 313
Qy 302 cggctccgcttgcttactgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 361
Db 314 GG-----GCTTATTGGGGCCAAAGGACCTGCTCACCCTCTCCTCAGGTGGAGCG 364
Qy 362 gttcagcggagtggtgctctgctgctgctgctgctgctgctgctgctgctgctgctg 421
Db 365 GTTCAGGCGGAGGTGGCTCTGCGGTGGCGGATCCCAATGTTCTCACCCAGTCTCCAG 424
Qy 422 caatcatgctctgcatctccaggggagagagagagagagagagagagagagagagagag 481
Db 425 CAATCATGCTGCATCTCCAGGGGAGAAAGTCAACATGACCTGCAGTGCACAGTTCAAGTG 484
Qy 482 taagtacatgcactggtaccagcagcagcagcagcagcagcagcagcagcagcagcag 541
Db 485 TAAGGTACATGAACCTGGTTCCAACAGAGTCAAGGACCTCCCGGAGGCTGGACATATG 544
Qy 542 acacatcaaaactgctctgagtcctgctgctgctgctgctgctgctgctgctgctgct 601
Db 545 ACACATCAAAACTGCTCTGAGTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
Qy 602 cttattctcacaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 661
Db 605 CTTACTCTCTCAATCAATCAGCAGCAGTGGAGGCTGGAAGTGTGCCACTTATTACTGCCAGC 664
Qy 662 agcagagtagttaccgctcagcagcagcagcagcagcagcagcagcagcagcagcagc 717
Db 665 AGTGGAGTAGTAATCACTCACTTTTCGGTGTGGAGCAACAGCTTGAACTGAACCG 720

RESULT 10
US-08-190-199A-64
; Sequence 64, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCHOV, Guy
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.

TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS

| | | | |
|----|-----|--|-----|
| Qy | 422 | caatcatgctcgtcatctccaaaggaggaaggtcaacatgacctgcagtgccagctcaagta | 481 |
| | | | |
| | | | |
| Db | 413 | CAATCATGTCTGATCTCCAGGGGAGAAGGTACCATGACCTGCAGTGCAGCTTCAAGTG | 472 |
| | | | |
| | | | |
| Qy | 482 | taagtatactgactggttacccagcagaagcgtgtaacctcccccaaaagatgattatg | 541 |
| | | | |
| | | | |
| Db | 473 | TAAGGTACATGACTGGTTCACACAGAAGTCAGGCACCTCCCCCAAAAGATGGATTATG | 532 |
| | | | |
| | | | |
| Qy | 542 | acacatccaaactggctcttgaggctccctgctcgtctgcttcagtgggcagtgggctgggaacct | 601 |
| | | | |
| | | | |
| Db | 533 | ACACATCCAAACTGTCTTGTGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCT | 592 |
| | | | |
| | | | |
| Qy | 602 | cttattctctcaaatcagcatgagaggtgtagatctccacttattactccatc | 661 |
| | | | |
| | | | |
| Db | 593 | CTTACTCTCTCAAACTACAGCAGCATGGAGGCTGAAGATGCTGCCATTTATTACTGCCACG | 652 |
| | | | |
| | | | |
| Qy | 662 | agcggagtagttaccgcgctcacgtctcggtgctgggacacagttggaaataaaacg | 716 |
| | | | |
| | | | |
| Db | 653 | AGTGAGGTAGTATTCACCTACCGTTCGGTCTGGGACCAAGCTGGAGCTGAACGG | 707 |
| | | | |
| | | | |

RESULT 11

```

US-08-190-199A-60
; Sequence 60, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCHOV, Guy T.
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.
; TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918

```

Query Match 55.4%; Score 397; DB 3; Length 711;
Best Local Similarity 73.3%; Pred. NO. 2e-103;
Matches 524; Conservative 0; Mismatches 182; Indels 9; Gaps 1;

| | | |
|-----|---|-----|
| 2 | agggtgaaactgcagcagctcaggagcctgaactggtgtnagcctgggggttcagtgaaagatat | 61 |
| 2 | ATGTGCAGCTGTGTGAGTCTGGGGAGGCTTAGTGCAGCTGTGAGGCTCCGGAACCTCT | 61 |
| 62 | cctgcgaagactcttganaaaaattcactgaatacacaccatgcactgggtgaagcagagcc | 121 |
| 62 | CTGTGTCAGCCCTCTGGATTCTACTTTCAGTAGCTTTGGAATGCCTGGGTTCGCAGGCCT | 121 |
| 122 | atggaaagagccttgagtgaggtatggaggtattaactcctaacaatggtggtactaaactaca | 181 |
| 122 | CAGAGAAGGGCTGGAGTGGGTCCATATATTAGTAGTGGCAGTAGTACCATCTACTATG | 181 |
| 182 | agcagaagtccaagggcgaagccattgactgtagcaagtcctccagcacagcctaca | 241 |
| 182 | CAGACACAGTGAAGGGCGGATTACCATCTCCAGAGACAATCCAGAACACCCCTGTTC | 241 |
| 242 | tggagctccgagcctgcacactgagagattctgcagctctattactgtgcaagagatacta | 301 |
| 242 | TGCAAAATGACCAGCTTAAGGTCTGAGGACAGGGCCATGTATTACTGTGCAAGAGATTACG | 301 |
| 302 | cgggtcccggttgccttactgggtccaagggacacaggtctcctcctcaggtggagcg | 361 |
| 302 | GG-----GCATTATTTGGGGCCAAAGGACTCTGGTCACTGCTCTGCAAGTGTGGTG | 352 |
| 362 | gttcaggcgaggtggctctgcggtggcgagatcggcacatcgagctcaactcagctctccag | 421 |
| 353 | GTAGCGGTGGTGGCGCAGTGGCGCGGCTCTCAAAATGTTCTTCCACAGCTCTCCAG | 412 |


```

US-08-235-838-15
: Sequence 15, Application US/08235838
: Patent No. 5571894
: GENERAL INFORMATION:
: APPLICANT: Wels, Winfried S.
: APPLICANT: Hynes, Nancy E.
: APPLICANT: Harwerth, Ina-Maria
: APPLICANT: Groner, Bernd
: APPLICANT: Hardman, NO. 5571894man
: APPLICANT: Zwackl, Markus
: TITLE OF INVENTION: Recombinant Antibodies Specific for a
: GROWTH FACTOR RECEPTOR
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CIBA-GEIGY Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: New York
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/235,838
: FILING DATE: TBA
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/828,832
: FILING DATE: 31-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 91-810079.3
: FILING DATE: 05-FEB-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919)541-8614
: TELEFAX: (919)541-8689
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2012 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Mouse/Pseudomonas aeruginosa

```

```
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pW215-51
FEATURE:
  NAME/KEY: CDS
  LOCATION: 1..1911
  OTHER INFORMATION: /note= "64-87 FLAG peptide and
  enterokinase cleavage site;97-456 FWP51 heavy
  chain variable domain;457-501 15 aa linker
  OTHER INFORMATION: sequence;502-822 FWP51 light chain variable
  OTHER INFORMATION: domain
  FEATURE:
  NAME/KEY: sig_peptide
  LOCATION: 1..63
  OTHER INFORMATION: /note= "ompA signal peptide"
  FEATURE:
  NAME/KEY: mat_peptide
  LOCATION: 94..1911
  FEATURE:
  NAME/KEY: 3'UTR
  LOCATION: 1912..2012
  OTHER INFORMATION: /function= "3' non-coding region of
  the exotoxin A gene"
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 825..1911
  OTHER INFORMATION: /note= "Exotoxin A gene coding
  region (coding for amino acids 252 to 613 of the
  mature exotoxin A)"
US-08-235-838-15
```

```
Query Match          48.6%; Score 348.8; DB 1; Length 2012;
Best Local Similarity 70.5%; Pred. No. 1.2e-89;
Matches 512; Conservative 0; Mismatches 199; Indels 15; Gaps 3;

QY 1  caggtgaactcagcagctcagcagcctgaactggtgagcctgggcttcagtgaaagata 60
DB 97  caggtgaactcagcagcagctcagcagcctgagcctgggcttcagtgaaagctg 156
QY 61  tcctgaagactcttggaacaaataactgaatacacaccatgactggtggaagcagagc 120
DB 157  tcctgaagactcttggaacaaataactgaatacacaccatgactggtggaagcagagc 216
QY 121  catggaagagccttgagtgagtgattgagtgatttaactcctaacaatggtggtactaac 180
DB 217  cctggacaagcccttgaatgattgattgattgattgattgattgattgattgattgattg 276
QY 181  aagcagaagttcaaggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
DB 277  aatcaaatgttcaaggacaagcagcagcagcagcagcagcagcagcagcagcagcagc 336
QY 241  atgagctccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
DB 337  atgcaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 396
QY 301  acggtccgc-----ttgttactggtggtccaaaggacacagcagcagcagcagcagc 351
DB 397  gcctctggggactggttctcgatgctggtggcccaaggacacagcagcagcagcagcagc 456
QY 352  ggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 411
DB 457  ggcgggtggcgggttctggtggtgggtggtggtggtggtggtggtggtggtggtggtggt 516
QY 412  cagtcctccagcaatcatgctgcatctccagggagagaggtcaccatgacgtcagtggc 471
DB 517  cagtcctccagcaatcatgctgcatctccagggagagaggtcaccatgacgtcagtggc 576
QY 472  agctca---agataagttacatgactggttacacagcagagcagcagcagcagcagcagc 528
DB 577  agccaagacacattaaagagttatagcttgggtaccacacacagcagcagcagcagcagc 636
QY 529  agatggatttatgacacacacacactggttctggtgagtcctcctgctcagtggcagtg 588
```

```
Db 637  CTACTCATACACTACACATCTGTATTACAGCCAGGATCCCATCCAGGTTTCAGTGGAACT 896
QY 589  gggctctgggagagattattcttctcacaatacagcagcagcagcagcagcagcagcagc 648
Db 697  GGGTCTGGGAGAGATTATTCTTTCAGCATCCACAACCTGGAGCTGAAGATATTGCAACT 756
QY 649  tattactgccatcagcggagtagtattaccgcgctcacgcttcgggtgctgggacacagttggaa 708
Db 757  TATTATG---TCTACATTATGATTATCTGTACACGTTCCGGAGGGGCGACCAAGCTGGAG 813
QY 709  ataaaa 714
Db 814  ATCAAA 819
```

Search completed: November 30, 1999, 14:10:18
Job time: 5144 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 12:44:33 ; Search time 545.98 Seconds
(without alignments)
5615.269 Million cell updates/sec

Title: US-08-940-544-3

Perfect score: 717

Sequence: 1 cagggtgaactgcagcagtc.....cacagttggaataaaccg 717

Scoring table: IDENTITY_NUC

Searched: 780561 seqs, 2137953050 residues

base :

GenEmbl.:

1: gb_bal:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pil:*
8: gb_pil2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_htg:*
19: em_hum1:*
20: em_hum2:*
21: em_in:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: gb_in3:*
37: gb_in4:*
38: gb_in5:*
39: gb_in6:*
40: gb_in7:*
41: gb_in8:*
42: gb_in9:*
43: gb_in10:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|----|----|-------------|
| ----- | | | | | | | |

| | | | | | | |
|----|-------|------|------|----|-----------|---------------------|
| 1 | 590.4 | 82.3 | 723 | 12 | AF035617 | AF035617 Mus muscu |
| 2 | 522.2 | 72.8 | 726 | 5 | A45598 | A45598 Sequence 29 |
| 3 | 521.6 | 72.7 | 726 | 5 | A45594 | A45594 Sequence 25 |
| 4 | 521.6 | 72.7 | 726 | 5 | AR063195 | AR063195 Sequence |
| 5 | 520.4 | 72.6 | 732 | 5 | E13410 | E13410 cDNA encodi |
| 6 | 517.8 | 72.2 | 2364 | 5 | E13412 | E13412 cDNA encodi |
| 7 | 517.4 | 72.2 | 1797 | 5 | I84705 | I84705 Sequence 2 |
| 8 | 512.6 | 71.5 | 726 | 5 | A45596 | A45596 Sequence 27 |
| 9 | 512.6 | 71.5 | 726 | 5 | AR063196 | AR063196 Sequence |
| 10 | 511.8 | 71.4 | 717 | 12 | AF025535 | AF025535 Mus muscu |
| 11 | 508 | 70.9 | 723 | 14 | AF169027 | AF169027 Synthetic |
| 12 | 507.2 | 70.7 | 810 | 5 | A45006 | A45006 Sequence 1 |
| 13 | 505.2 | 70.5 | 726 | 12 | MMU20617 | U20617 Mus musculu |
| 14 | 504.6 | 70.4 | 732 | 5 | A45588 | A45588 Sequence 19 |
| 15 | 504.6 | 70.4 | 732 | 5 | A45600 | A45600 Sequence 31 |
| 16 | 504.6 | 70.4 | 732 | 5 | AR063192 | AR063192 Sequence |
| 17 | 503 | 70.2 | 732 | 5 | A45590 | A45590 Sequence 21 |
| 18 | 503 | 70.2 | 732 | 5 | AR063193 | AR063193 Sequence |
| 19 | 502.8 | 70.1 | 738 | 12 | AF074900 | AF074900 Mus muscu |
| 20 | 501.6 | 70.0 | 1679 | 5 | AR054409 | AR054409 Sequence |
| 21 | 493.6 | 68.8 | 798 | 12 | AF000955 | AF000955 Mus muscu |
| 22 | 487 | 67.9 | 738 | 5 | A45592 | A45592 Sequence 23 |
| 23 | 487 | 67.9 | 738 | 5 | AR063194 | AR063194 Sequence |
| 24 | 456.6 | 63.7 | 1944 | 14 | XXU78761 | U78761 Synthetic c |
| 25 | 449.2 | 62.6 | 678 | 12 | AF003714 | AF003714 Mus muscu |
| 26 | 445.4 | 62.1 | 596 | 12 | AF003724 | AF003724 Mus muscu |
| 27 | 438 | 61.1 | 684 | 12 | AF003720 | AF003720 Mus muscu |
| 28 | 434.8 | 60.6 | 696 | 12 | AF003721 | AF003721 Mus muscu |
| 29 | 429.6 | 59.9 | 687 | 12 | AF003725 | AF003725 Mus muscu |
| 30 | 421.6 | 58.8 | 702 | 12 | AF003708 | AF003708 Mus muscu |
| 31 | 418.6 | 58.4 | 738 | 12 | AF141321 | AF141321 Mus muscu |
| 32 | 414.2 | 57.8 | 729 | 12 | AF002242 | AF002242 Mus muscu |
| 33 | 413.2 | 57.6 | 738 | 12 | MMU40581 | U40581 Mus musculu |
| 34 | 412.4 | 57.5 | 720 | 12 | MMU40582 | U40582 Mus musculu |
| 35 | 412.4 | 57.5 | 4354 | 14 | ASY14583 | Y14583 Artificial |
| 36 | 410.8 | 57.3 | 738 | 5 | A41674 | A41674 Sequence 1 |
| 37 | 403.4 | 56.7 | 684 | 12 | AF003711 | AF003711 Mus muscu |
| 38 | 397 | 55.4 | 711 | 5 | A25748 | A25748 Variable re |
| 39 | 397 | 55.4 | 711 | 5 | AR051416 | AR051416 Sequence |
| 40 | 391.4 | 54.6 | 783 | 12 | AB007986 | AB007986 Mus muscu |
| 41 | 390.6 | 54.5 | 1314 | 5 | A68604 | A68604 Sequence 4 |
| 42 | 390 | 54.4 | 684 | 12 | AF003716 | AF003716 Mus muscu |
| 43 | 388.6 | 54.2 | 783 | 12 | MMU010065 | MMU010065 Mus muscu |
| 44 | 386.8 | 53.9 | 708 | 5 | A25744 | A25744 Variable re |
| 45 | 386.8 | 53.9 | 708 | 5 | AR051414 | AR051414 Sequence |

ALIGNMENTS

| RESULT | 1 | AF035617 | Mus musculus scFv antibody SS mRNA, partial cds. | 723 bp | AF035617 | 30-SEP-1999 |
|------------|---|--|--|----------|-------------|-------------|
| LOCUS | AF035617 | Mus musculus scFv antibody SS mRNA, partial cds. | 723 bp | AF035617 | 30-SEP-1999 | |
| DEFINITION | AF035617 | Mus musculus scFv antibody SS mRNA, partial cds. | 723 bp | AF035617 | 30-SEP-1999 | |
| ACCESSION | AF035617 | Mus musculus scFv antibody SS mRNA, partial cds. | 723 bp | AF035617 | 30-SEP-1999 | |
| NID | 92921393 | GI:2921393 | | | | |
| VERSION | AF035617.1 | GI:2921393 | | | | |
| KEYWORDS | | | | | | |
| SOURCE | house mouse. | | | | | |
| ORGANISM | Mus musculus | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. | | | | | |
| AUTHORS | Chowdhury, P.S., Viner, J.L., Beers, R. and Pastan, I. | | | | | |
| TITLE | Isolation of a high-affinity stable single-chain Fv specific for mesothelin from DNA-immunized mice by phage display and construction of a recombinant immunotoxin with anti-tumor activity | | | | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 95 (2), 669-674 (1998) | | | | | |
| MEDLINE | 98118570 | | | | | |
| REFERENCE | 2 (bases 1 to 723) | | | | | |
| AUTHORS | Chowdhury, P.S. | | | | | |
| TITLE | Direct Submission | | | | | |
| JOURNAL | Submitted (24-NOV-1997) LMB, NCI, 37 Convent Drive, Bld. 37 Rm. | | | | | |

FEATURES
source

4B20, Bethesda, MD 20892-4255, USA
Location/Qualifiers
1..723
/organism="Mus musculus"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="spleen"
/clone_lib="phage display library made from spleen mRNA of mice immunized with DNA"
1..>723
CDS
1..>723
/codon_start=1
/product="scFv antibody SS"
/protein_id="AAC04760.1"
/db_xref="PID:g2921394"
/db_xref="GI:2921394"
/translation="MQVLOQSGPELEKPKGASVKISKASKSYSTGYTMNWKOSHCK SLEWGLITPYNGASSYNOKRGRATLVDKSSYAYMDLLSLTSDSAVFCAARGY DARGDYWGOCCTTVNSGVGGSGGGSDIELTQSPAINASFGKRYHTCSA SSVSYMWOYOKTSFKPRMIDYTKLASGVPRFSGSGSNYSITISVFAEDDA TTYCWMSGYPLTFAGATKLEIK"
176 a 180 c 203 g 164 t
ORIGIN

Query Match 82.3%; Score 590.4; DB 12; Length 723;
Best Local Similarity 89.7%; Pred. No. 7.5e-172;
Matches 646; Conservative 0; Mismatches 68; Indels 6; Gaps 1;

QY 1 cagtgaaactcagcagtcaggacctaagctggtgnagcctgggccttcagtgaagata 60
|||||
Db 4 CAGGTACACTGCAGCAGCTGTGGCGCTGAGCTGGAGAGCCGTGGCGCTTCAGTAGAATA 63
|||||

QY 61 tcctgcaagactcttggnacaaattcactgaatcacaccatcactggtgtaagcagagc 120
|||||
Db 64 TCCTGCAAGCCTTCTGTTACTCATTTACTTGCTGCTACACCATTGCTGGTAGCAGAGC 123
|||||

QY 121 catggaagaagcccttgatgtagttgaggtatataacctacaacatggtggtactaacac 180
|||||
Db 124 CATGGAAGAAGCCTTGAGTGGATTGGACTTATTACTCTTACAATGGTCTTACTAGTAC 183
|||||

QY 181 aagcagaagtcgaagggcgaagcccatgactgctagacaaagctccacagcagcctac 240
|||||
Db 184 AACCAAGAATTGAGGGGCAAGGCCATTAATCTGTAGACAAGTATCCAGCACAGCCTAC 243
|||||

QY 241 atggagctccgcagctgacatctgaggattctgcagttattactgtgcaagagatact 300
|||||
Db 244 ATGGAGCTCTCAGCTGACATCTGAAGACTCTGCAGTCTATTCTGTGCAAGGGGGGT 303
|||||

301 acggtccc-----ttcttactggtgccaaaggaacagctcaccgtctctcaggt 354
|||
Db 304 TAGCAGCGGAGGGGTTTTGACTACTTGGGCCCAAGGAGCACCGTCCACCGTCTCTCAGGT 363
|||||

QY 355 ggaggcggttcaggcgaggtgctctgctggtgctggaatcgacatcagctcactcag 414
|||||
Db 364 GRAGGCGGTTACGGCGCGGTGGCTCTGCGGTGGCGATCGGATCGAGCTCAGCTCAG 423
|||||

QY 415 tctccagcaatcatgtctgcatctccaggggagaaggtcaccatgacctgcagtggcagc 474
|||||
Db 424 TCTCCAGCAATCATGCTCTCATCTCCAGGGGAGAGGTCACCATGACCTGCGAGTCCAGC 483
|||||

QY 475 tcaagataagttacatgcactggttaccagcagaagcctgtcaccctccccaaaagatgg 534
|||||
Db 484 TCAAGTGTAAATTATATGCTGTTGGTACCAGCAGAAGTCAGGCACCTCCCCAAAAGATGG 543
|||||

QY 535 atttatgacacatccaactgctcttgagtcctctgctcagtgagtgagtggtctct 594
|||||
Db 544 ATTATGACACATCCAACATGGCTCTGAGGTCCAGAGTCCAGTCCAGTGGCAGTGGGTCT 603
|||||

QY 595 gggaacctcttctctcaaatcagcagcagcagcagcagcagcagcagcagcagcagcag 654
|||||
Db 604 GGAACTCTTACTCTCATATCAGCAGCTGGAGGCTGAAGATGATGATGATGATGATGAT 663
|||||

QY 655 tgccatcagcgagtagttaccocgcctcagttcgtgctgggacacagcttggaataaaa 714
|||||

RESULTS
A46598 726 bp DNA PAT 07-MAR-1997
LOCUS Sequence 29 from Patent WO9525167.
DEFINITION A46598
ACCESSION 92300763
NID A46598.1 GI:2300763
VERSION house mouse.
KEYWORDS Mus musculus
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Other publication ZA 9502174 951227
Other publication CN 1124501 960612
Other publication HU 73461 960828
Other publication CZ 9503014 960214
Other publication PL 311661 960304
Other publication CA 2163012 950921
Other publication NO 954626 951116
Other publication AU 2071695 951003.
Location/Qualifiers
1..726
/organism="Mus musculus"
/strain="BALB/C"
/db_xref="taxon:10090"
/dev_stage="ADULT"
/tissue_type="SPLENOCYTES"
1..726
/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02932.1"
/db_xref="PID:e306394"
/db_xref="PID:g2300764"
/db_xref="GI:2300764"
/translation="EVKLQSGAEVLKPGASVKLSCKASYFTSHMHVKKAKORGL EWGEFNPSNGRTNYNEKFASKATLVLDKSSRAYMLQLSSLISEDASVYICSRDIL YDRIFDYMGOGTIVTVSSGGGGGGGGSDIELTQSPIMSASPGEKVMTICS DSSVSVMYWOQQKTGSSPRLIYDTNLASGVPRFSGSGSTSYSLTISRMEADA ATYCQOWSSYPHFGATKLEIK"
173 a 192 c 206 g 155 t
BASE COUNT
ORIGIN

Query Match 72.8%; Score 522.2; DB 5; Length 726;
Best Local Similarity 84.6%; Pred. No. 8.8e-151;
Matches 613; Conservative 0; Mismatches 100; Indels 12; Gaps 2;

QY 2 aggtgaaactcagcagtcaggacctaagctggtgnagcctgggccttcagtgaagatat 61
|||||
Db 2 AGGTCAAGCTGCAGCAGCTCAGGGGCTGAAGCTGGTGAAGCTGGGCTTCAGTGAAGTTGT 61
|||||

QY 62 ctgcgaagactctggnacaaattcactgaatacacatcactggtgtaagcagagcc 121
|||||
Db 62 CTGCAAGGCTTCCCGCTACACCTTCACAGCCACTGGATGACATGGGTGAAGCAGGG 121
|||||

QY 122 atggaagagccttgagtgattgaggtattaatcctaacaatggtggtactaactaca 181
|||||
Db 122 CTGGCAAGCGTGGAGTGGATCGGAGAGTTAATCCCAGCAACGCCGCTACTAATAC 181
|||||

QY 182 agcagaagttcaaggcgaagccacattgactgtgatacaagtctccacagcagcctaca 241
|||||
Db 182 ATGAGAAATTCAGAGCAGGACCACTGACTGTAGACAAATCTCTCCAGCAGCAGCTACA 241
|||||

```
QY 242 tgaagctccgagcctgacatctgaggtattctgagttctattactgtgcaag-agataact 300
Db 242 TGCAACTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCCAGTCGGGACT 301
QY 301 acgggtcccg-----tttgcttactgggtccaaggggaccaggtccacgtctcct 349
Db 302 ATGATTACGACGGACGGTACTTTGACTACTGGGGCCCAAGGGACACGGTCACCGTCTCCT 361
QY 350 caggtggagggcgttcaggcgaggtggtctctggcggtggcggtgagatcgacatcgagctca 409
Db 362 CAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGCGGGATCTGACATTGAGCTCA 421
QY 410 ctcaattccagcaatcatgtctgcatctccaaggggagaaggtcaaccatgacctgcagt 469
Db 422 CCCAGTCTCCATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 470 cgagctcaagtaagttacatgctgctggtaccagcagaagcctgtcacctcccccacaaa 529
Db 482 ACAGCTCAAGTGTAGTTACATGTACTGGTACACAGACAGACAGATCCCTCCCCAGAC 541
QY 530 gatgatttatgacatcccaactgggtcttggtgagtcctctgctgctcagtggtgagtg 589
Db 542 TCCTCATTTATGACACATCCAACTGGCTCTGGAGTCCCTGCTTCAGTGGCAGTG 601
QY 590 ggtctgggacctcttattctctcaaatcagcagcatgaggtgagctgtagatgctgcaatt 649
Db 602 GGTCTGGGACCTTCTACTCTCTCAATCAGCCGAATGAGGCTGAAGATGTGCCACTT 661
QY 650 attactgcatcagcaggaagtaagtcacgcctcagctcagcttgcgtctgggacagattgaaa 709
Db 662 ATTACTGCCAGCAGTGGAGTAGTTACCCACACACACACCTGCGTCTGGGACCAAGCTGGAAA 721
QY 710 taaaa 714
Db 722 TAAAA 726

RESULT 3
LOCUS A46594 726 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 25 from Patent WO9525167.
ACCESSION A46594
NID 92300759
VERSION A46594.1 GI:2300759
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 726)
AUTHORS Kettleborough,A.C., Bendig,M.M., Ansell,K.H., Guessow,D., Adan,J.,
Mitjans,F., Rosell,E., Blasco,F. and Pluats,J.
TITLE ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR ANTIBODIES
JOURNAL Patent: WO 9525167-A 25 21-SEP-1995;
MERCK PATENT GMBH (DE)
COMMENT Other publication ZA 9502174 951227
Other publication CN 1124501 960612
Other publication HU 73461 960828
Other publication CZ 9503014 960214
Other publication PL 311661 960304
Other publication CA 2163012 950921
Other publication NO 954626 951116
Other publication AU 2071695 951003.
Location/Qualifiers
1..726
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/tissue_type="SPLENOCYTES"
1..726
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"

CDS
```

```
/codon_start=1
/protein_id="CAA02930.1"
/db_xref="PID:e306173"
/db_xref="PID:g2300760"
/db_xref="GI:2300760"
/translation="QVKLOESGAELVKPGASVLSCKASGYTFTSHMHWKORAGOG
LEWTEINPTAPNTNNEKPKATLTVDKSSYAYMOLSSLSESAVYXASRDYD
YDRYRDYWGCGTIVTVSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
DSSSVSYWYQKTSPPRLLIYDTSNLASGVFVRFGSGSGSTSYSLTISRMEADA
ATYCCOWSSYPLTFGAGTKLEIK"
BASE COUNT 174 a 192 c 206 g 154 t
ORIGIN
```

```
Query Match 72.7%; Score 521.6; DB 5; Length 726;
Best Local Similarity 84.4%; Pred. No. 1.4e-150;
Matches 613; Conservative 0; Mismatches 101; Indels 12; Gaps 2;

QY 1 caggtgaaactgcagcagctcaggaacctgagcctggtgagcctggggttcagtaagata 60
Db 1 CAGGTGAAACTGCAAGGAGTCTGGGGCTGAAGTGTGAGGCTGGGGCTTCAGTGAAGTTG 60
QY 61 tctgtcaagactctgaaacaaatttcactgaatacacacatgcactgggtgaagcagagc 120
Db 61 TCTGTCAAGGCTTCGGGCTACACCTTCACGAGCCACTGGATGCACCTGGGTGAAGCAGAGG 120
QY 121 catgaaagagccttgcagtgagtgattggaggtattaatcctaaacaatggtggtactaac 180
Db 121 GCTGCACAAGGCCCTTGAGTGGATCGGAGATTAAATCCAGAACGGCGCTTACTTAAC 180
QY 181 aagcagaagttcaaggcgaagccacattgactgtagaagtcctccagcagcagctac 240
Db 181 AATGAGAAATTCAGAGACAGAGCCACACTGACTGTAGACAAATCTCCAGCACAGCCTAC 240
QY 241 atgagctccgcagcctgacatctgaggtattctcagcttacttactgtgaag-agatac 299
Db 241 ATGCAACTCAGCAGCCTGACATCTGAGGACTCTCGGGTCTATTACTGTGCGAGTCGGGAC 300
QY 300 tacgggtcccg-----tttgcttactgggtccaaaggagaccaggtcacggtctcc 348
Db 301 TATGATTACGACGGCAGGTACTTTGACTACTGTGGGCGCAAGGACACAGGTCACCGTCTCC 360
QY 349 tcaggtgagcgggttcaggcggagg-ggctctgctggcgggtggcgatcgagcatcgagctc 408
Db 361 TCAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 420
QY 409 actcagttccagcaatcatgtctgcattccaggggagaggtcaccatgacctgcagt 468
Db 421 ACCAGTCTCCCAACAATCATGTCTGCATCTCCAGGGAGAGAGGTACCATGACCTGCAGT 480
QY 469 ggcagctcaagataaagttacatgcactggtaccagcagaagcctgtcacctcccccaaa 528
Db 481 GACAGCTCAAGTGTAGTTACACCTACTGTTACAGCAGAGACAGGATCTCTCCCCCAGA 540
QY 529 agatggatttatgacacatccaaactggctctctggagtcctctgctgcttcaagtgagct 588
Db 541 CTCCTGATTTATGACACATCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 589 gggcttgggaccttattctcacaatcagcagcatggaggtgtagatgtgcacct 648
Db 601 GGGCTGGGACCTTCTTACTCTCTCACAATCAGCGAATGGAGGCTGAAGATGTGCCACT 660
QY 649 tattactgcatcagcaggtagttaccgcgtcacgttcggtgctgggacacagttgaa 708
Db 661 TATTACTGCCAGCAGTGGAGTAGTTACCCGCTCAGTTCGTTGGTGGTGGTGGTGGTGGTGG 720
QY 709 ataaaa 714
Db 721 ATAAAA 726

RESULT 4
AR063195
```

[illegible]


```
Qy 536 ttatgacacatcaaaactggtcttctgagtcggtccctgctgcttcaagtggcagtggtctg 595
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 545 TTTATGGCACATCAAACTGGCTTCTGGAATCCCTGTTTCGTTTCAGTGGCAGTGATCTG 604
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 596 ggacctctattctctcaaatcagcagcatggagctgtagatgctgccattattact 655
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 605 GGACCTCTATTCTCTACATCAGCAGCATGGAGGGTGAAGATGCTGCCACTTATTACT 664
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 656 gccatcagcggagtagttaccgcctcacgttctcggtgctggtggacacagttggaataaac 715
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 665 GCCAGCAGTGGAGTAGTAACCATTCACGTTCTCGGCTCGGGGACCAAGCTGGAATATAAG 724
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 716 g 716
Db 725 G 725

RESULT 7
INIS 184705 184705 1797 bp DNA PAT 04-APR-1998
DEFINITION Sequence 2 from patent US 5696237.
ACCESSION 184705
NID 93022225
VERSION 184705.1 GI:30222225
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1797)
AUTHORS Fitzgerald, D., Chaudhary, V. Kumar, Pastan, I. Harry,
Waldmann, T. Alexander and Queen, C.L.
TITLE Recombinant antibody-toxin fusion protein
JOURNAL Patent: US 5696237-A 2 09-DEC-1997;
FEATURES Location/Qualifiers
source 1. 1797
BASE COUNT 336 a 598 c 571 g 292 t
ORIGIN /organism="unknown"

Qy 1 caggtgaaactgcagcagctcaggacctgaactggtgagcctgggtcttcagtgaaagata 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4 CAGGTCCAGCTGCAGCAGCTCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATG 63
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
61 tccctgcaagacttctgganacaaattcactgaataacacatgcactggtggaagcagagc 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
64 TCCTGCAAGGCTTCTGGCTACACCTTTACTAGCTACAGGATGCAGTGGTAAAAACAGAGG 123
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 121 catggaagagccttgatgagtgaggtattaatcctcaaatggtggttactaactac 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 124 CCTGGACAGGCTCGGATGATGGATATATTAATCCTACAGCTGGTATCTACTCAATAC 183
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 181 aagcagaagtccaagggcgaagccacattgactgtagaacaagtctccagcagacgtcac 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 184 AATCAGAAGTTCAAGGACAAGGCACATTTGACTGCAGACAAATCTCCAGCACAGCCTAC 243
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 241 atgagctcccgagcctgacatctgaggtattctcagcttattactgtgaagagatact 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 244 ATGCAACTGACGACCTGACATTTGAGGACTCTGCAGTCTATTACTGTGCAAGAGGGGGG 303
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 301 acggtcccgcttctgactggttccaaaggaccacggtccacgctctctcactcaggtggaggc 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 304 GGGGTC---TTTGACTACTGGGGCCAGGAACCACTCTCAGCTCTCTCCGGAGGGCGGT 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 361 ggttcagcgaggtggtctcgtggtggtggtggtggtggtggtggtggtggtggtggtggt 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 GGTCTGGGCGGTGGGGCTCGGGTGGGGCGGCTCTCAAAATGTTCTCACCCAGCTCTCCA 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 421 gcaatcagctgcatctcctcaggggaggaaggtcaccatgacactgcagtgagcagctcaagt 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```
Db 421 GCAATCATGCTGCATCTCCAGGGGAGGAAGGTCAACATAACCTGCAGTGCAGCTCAAGT 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 481 ataagttacatgcactggtaccagcagaagccctgtcacctcccccaaaagatgattat 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 481 ATAAGTTACATGACTGCTTCCAGCAGAACCCAGGCACTTCTCCCAAACTCTGGATTAT 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 541 gacacatcccaaaactggcttctgagtcctcctgctcgtctcagtcagtgagtggtctgggacc 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 541 ACCACATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGCAGTGCAGTGGGACC 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 601 tcttattctctcaaatcagcagcatggagctgtagatgctgccattattactgcat 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 601 TCTTACTCTCTCAATCAGCGAATGGAGGCTGAAGATGCTGCCACTTATTACTGCCAT 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 661 cagcggagtagttaccgcctcacgttctcggtgctggtggacacacagttggaataaa 713
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 661 CAAGGAGTACTTACCACCTACGTTCTGGTTCGGGACCAAGCTGGAGCTCAA 713
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 8
LOCUS A46596 726 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 27 from Patent WO9525167.
ACCESSION A46596
NID 92300761
VERSION A46596.1 GI:2300761
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 726)
AUTHORS Kettleborough, A.C., Bendig, M.M., Ansell, K.H., Guessow, D., Adan, J.,
Mitjans, F., Rosell, E., Blasco, F. and Piuels, J.
TITLE ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR ANTIBODIES
JOURNAL Patent: WO 9525167-A 27 21-SEP-1995;
COMMENT MERCK PATENT GMBH (DE)
Other publication ZA 9502174 951227
Other publication CN 1124501 960612
Other publication HU 73461 960828
Other publication CZ 9503014 960214
Other publication PL 311661 960304
Other publication CA 2163012 950921
Other publication NO 954626 951116
Other publication AU 2071695 951003.
FEATURES Location/Qualifiers
source 1. 726
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/tissue_type="SPLENOCYTES"
/clone_lib="7 G 1 (SINGLE-CHAIN FV, HEAVY, LIGHT CHAIN,"
/clone_lib="LINKER)"
1. 726
/notes="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02931.1"
/db_xref="PID:8306174"
/db_xref="PID:92300762"
/db_xref="GI:2300762"

CDS
171 a 193 c 205 g 157 t
BASE COUNT
ORIGIN
```

Query Match 71.5%; Score 512.6; DB 5; Length 726;
Best Local Similarity 83.7%; Pred. No. 8.2e-148;

| | | | | | | | | | |
|---------|------|--------------|----|------------|------|--------|-----|------|----|
| Matches | 607; | Conservative | 0; | Mismatches | 106; | Indels | 12; | Gaps | 2; |
|---------|------|--------------|----|------------|------|--------|-----|------|----|

2 aggtgaactcagcagctcaggaccgaaactgctgnagcctgggggttcagtgaagatat 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db AGCTCAAGCTGCAGCAGCTAGGGGCTGAACGTGTGAAGCTGGGGCTTCAGTGAAGTTGT 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 cctgcagactctganacaaattcactgaatatcacaccatgcactgggtgaagcagagcc 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 CCATCAAGGCTTCGGCTCACCTTCACAGCCACTTGGATCTACTGGGTGAACAGCAGGG 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 atggaagagccttgagtgattggagggtattaatcctaacaattggtgtactaaactaca 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 GCTGGCAAGGCCTTGAGTGGATCGGACAGTTTAATCCCAACAAGCCGCTACTAATCTACA 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 agcagaagttcaaggccaaggcacattgactctagacaagtctccagcacagcctaca 241
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 ATGAGAAATTCAGAGCAAGGCCACACTGACTGTAGACAATCTCCAGCACAGCCTACA 241
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 242 tggagctccgagcctgacatctgaggatctcagctctattactgtgcaag-agatact 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 242 TCGAATCAGACGCTGACAATCTGAGGACTGCTGGGTCTATTACTGTGCCAGTCGGGACT 301
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 acggtcccg-----ttgtctactgggtcccaaggagcacccaggtctacccgtctcct 349
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 302 ATGATTACGAGCGAGGGTACTTTTGACTACTTGGGCCCAAGGACCACGGTCAACGTCCTCT 361
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 caggtggagcggttcaggcgaggttggtctctggcggtggcgatcggacatcgagcctca 409
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 362 CAGGTGGCGGTGCTCGGGCGGTTGGTGGTGGGCTGGCGGATCTGACATTGAGCTCA 421
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 410 ctcagtctccagcaatcatgtctgcattctccaggggaggaaggtccacatgacctcgagtg 469
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 422 CCAGTCTCCAACATCATGCTTCGATCTCCAGGGGAGAAGTTCACCATGACCTGGCAGTG 481
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 470 gcagctcaagtataaagttacatgcactggtaccagcagaagcctgtcacctcccccaaaa 529
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 482 ACAGCTCAAGTGAAGTTACATGACTGTGTACCAGCAGAAGACAGGATCTCTCCCCCAGAC 541
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 530 gatggattatgacacatccaaactggctctcctgagtgccctgctcgtctcaatggcagtg 589
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 542 TTCTGATTTATGACATCCAACTGGCTTCTGGAATCCCTGTCTGCTTTCAGTGGCAGTG 601
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 590 ggtctgggacctctattctctcacatcagcagcatggaggtgttagatgctgccacit 649
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 602 GGTCTGGGACCTCTTACTCTCTCACAAATCAGCCGAATGGAGGCTGAAGATGCTGCCACIT 661
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 650 attactgccatcagcgagtagttaccctgcactcgttcggtctgggtgggacacagttggaaa 709
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 662 ATTACTGCCAGCAGTGAGGTAGTTACCCGCTCAGTTCGGTGTGGGACCAAGCTGGAAA 721
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 710 taaaa 714
| | | | |
Db 722 TAAAA 726
| | | | |

| | |
|------------|--|
| RESULT | 9 |
| AR063196 | |
| LOCUS | 736 bp DNA |
| DEFINITION | Sequence 27 from patent US 5844093. |
| AR063196 | PAT |
| ACCESSION | |
| NID | 95990887 |
| VERSION | AR063196.1 GI:5990887 |
| KEYWORDS | . |
| SOURCE | Unknown. |
| ORGANISM | Unknown. |
| REFERENCE | Unclassified. |
| AUTHORS | 1. (bases 1 to 726) |
| TITLE | Kettieborough,A.Cathrine, Bendig,M.M., Ansell,K.H., Gussow,D. Adan,J., Mitjans,F., Rosell,E., Blasco,F. and Pilats,J. |
| JOURNAL | Anti-EGFR single-chain Fvs and anti-EGFR antibodies |
| FEATURES | Patent: US 5844093-A 27 01-DEC-1998; Location/Qualifiers 1..726 |

| BASE COUNT | 171 a | /organism="unknown" | 193 c | 205 g | 157 t |
|------------|-------|---------------------|-------|-------|-------|
| ORIGIN | | | | | |

Query Match 71.5%; Score 512.6; DB 5; Length 726;

| | | | |
|-----------------------|-------------------|---------------|-------------|
| Best Local Similarity | 83.7% | Fied: NO. | 8.2E-148; |
| Matches | 607; Conservative | 0; Mismatches | 106; Indels |
| | | 12; Gaps | 2; |

| | | | |
|----|-----|--|-----|
| Qy | 2 | agggtgaactcgagcagctcaggaccgaaactgctgnaagcctggggcttcagtgaaatat | 61 |
| Db | 2 | AGGTCAAGCTTCGACGAGTCAGGGGCTGAACCTGGTGAAGCCTGGGGTCTCAGTGAAGTTGT | 61 |
| Qy | 62 | cctgcagaactctcganacaaattccactgaatacacaccatgcactgaggtgaacgaagcc | 121 |
| Db | 62 | CCTGC AAGGCTTCGGGCTACACCTTCACCAAGCCACTTGGATCACTGGGTGAAGCAGAGGG | 121 |
| Qy | 122 | atggaagaagccttgatggatggagggtattaatcctaacaatggtggctactaaactaca | 181 |
| Db | 122 | GCTGGCAAGGCCCTGAGTGGATCGGACAGTTTAATCCCACGACAGGCCGTTACTTAACATA | 181 |
| Qy | 182 | agcagaagttcaagggcaaggccacatcttgactctagacaagttcctccoaagcacagctaca | 241 |
| Db | 182 | ATGAGAAATTC AAGAGCAAGCCGACACACTGACTGTAGACAAATCCTCGACGACAGCCTACA | 241 |
| Qy | 242 | tgaagctccgcagcctgacatctgaagattctcaagtctcaactctattactgtgaag-agatact | 300 |
| Db | 242 | TCGAACATCAGACGCTGACATCTGAGGACTGCTCGGTCTATTACTGTGCCAGTCGGGACT | 301 |
| Qy | 301 | acggtcccg-----tttgcttactgggtcccaaggagcacagggtcacccgtctcct | 349 |
| Db | 302 | ATGATTACGACGGACGGTACTTTTGACTACTGGGCCAAGGACCACGGTCAACCGTCTCCT | 361 |
| Qy | 350 | caggctggaagcgtgtcaaggcggaagggtggtctctgctggcgtggcgatcggacatcgagctca | 409 |
| Db | 362 | CAGGTGGCGTGGCTCGGGCGGTGGTGGGTGGGTGGCGGATCTGACATTGAGCTCA | 421 |
| Qy | 410 | ctcagctccagcaaatcatgtctgcattctccaggggagaaggtccacatgacctgcagtg | 469 |
| Db | 422 | CCCAGTCTCCAAACATCATGCTGCAATCTCCAGGGAGAGGTCACCATGACCTGCAGTG | 481 |
| Qy | 470 | gcagctcaagtataaagttacatgcactggtaaccagcagaagcgtgtcacctcccccaaa | 529 |
| Db | 482 | ACAGCTCAAGTGTAA GTTACATGTA CTACTGGTACCAGCAGAAGCAGGATCCTCCCCAGAC | 541 |
| Qy | 530 | gatggatttatgacacatccaaactggcttctggagtccctgctcgtctcagtyggcagtg | 589 |
| Db | 542 | TTCTGATTTATGACATCCAACTCGCTTCTGGAGTCCCTGTTCCGCTTCAGTGGCAGTG | 601 |
| Qy | 590 | ggtctgggaacctttattctctcacaaatcagcagaatggaagcgtgtagatgctgcacatt | 649 |
| Db | 602 | GGTCTGGACCTCTTACTCTCACAATCAGCCGAATGAGGCTGAAGATGTGTGCCACAT | 661 |
| Qy | 650 | attactgccaatcagcgagtagttaccgcgtcacgttcggtgctgggacacagttggaaa | 709 |
| Db | 662 | ATTACTGCCAGCAGTGGAGTAGTTACCCGCTACAGTTCGGTGTGGGACCAAGCTGGAAA | 721 |
| Qy | 710 | taaaa | 714 |
| Db | 722 | TA AAA | 726 |

RESULT 10

| | | | | | |
|------------|--|------------|------|-----|-------------|
| AF025535 | AF025535 | 717 bp | mRNA | ROD | 07-JUL-1998 |
| LOCUS | Mus musculus single chain Fv antibody (E4(Fv)) mRNA, partial cds | | | | |
| DEFINITION | AF025535 | | | | |
| ACCESSION | AF025535 | | | | |
| NID | 93293252 | | | | |
| VERSION | AF025535.1 | GI:3293252 | | | |
| KEYWORDS | | | | | |
| SOURCE | house mouse. | | | | |
| ORGANISM | Mus musculus | | | | |
| | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; | | | | |

[illegible]

| | |
|-----------------------|---|
| KEYWORDS | unidentified. |
| SOURCE | unidentified |
| ORGANISM | unclassified. |
| REFERENCE | 1 (bases 1 to 810) |
| AUTHORS | Chester, K.A., Hawkins, R.E. and Begent, R.H. |
| TITLE | ANTIBODY AGAINST CARCINOEMBRYONIC ANTIGEN (CEA) |
| JOURNAL | Patent: WO 951341-A 1 08-JUN-1995; |
| COMMENT | CANCER RES CAMPAIGN TECH (CB) |
| | Other publication CA 2177584 950608 |
| | Other publication AU 1194795 950619. |
| FEATURES | Location/Qualifiers |
| source | 1..810 |
| | /organism="unidentified" |
| | /db_xref="taxon:32644" |
| CDS | 1..810 |
| | /note="unnamed protein product; Protein sequence is in conflict with the conceptual translation" |
| | /codon_start=1 |
| | /protein_id="CAA02813.1" |
| | /db_xref="PID:e306097" |
| | /db_xref="PID:g2299568" |
| | /db_xref="GI:2299568" |
| | /translation="ETVIMKYLPTAAAGLLLLAAQPAMQVKLQSGAEVLRSVGRSTVS KLSTFASGNFKSDYHMLRWQPEGLWIDPENGDEYAPKFGKATFTDTSS ENVALQLSSLTSEDYAVYCNQETPTGYFDYWGQCTTVTVSGGGSGGGSGGGG SENVLTQSPATMSASPGKVTITCSASSSYXMHVFOOKPCTSPKLIWYSTSLNLSAG TPARSGSGSGTYSYLSITSRMEADDAATYYCQRRSSYPLTFGAGTKLEKRAA" |
| BASE COUNT | 189 a 220 c 222 g 179 t |
| ORIGIN | |
| Query Match | 70.7%; Score 507.2; DB 5; Length 810; |
| Best Local Similarity | 82.2%; Pred. No. 3.9e-146; |
| Matches | 597; Conservative 0; Mismatches 120; Indels 9; Gaps 1; |
| QY | 1 caggtgaaactcagcagtcaggacactggaactggctgnaagctgggcttcagtgaaagata 60 |
| Db | 79 CAGGTCAAACCTCAGCAGCTCTGGGCAGAACTTGTGAGGTCAGGACCTCAGTCAAGTTG 138 |
| QY | 61 tctgtcaagactcttgannacaaattcactgaatacacacacacacacacacacacagc 120 |
| Db | 139 TCTGTGACAGCTCTTGCTTCAACATTAAGACTCTTATGCACTGTTGTGAGCGAGGG 198 |
| QY | 121 catggaaagacgttgagtgagtgattgagtgatttaactcctaacaatgggtgactaacac 180 |
| Db | 199 CTTGAACAGGGCTGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 258 |
| QY | 181 aagcagaagtccaaggcagccacattgactgtagacaagtcctccagcacagcctac 240 |
| Db | 259 GCCTCCGAAGTTCCAGGCAAGCCACTTTACTACAGACACATCCTCCACACAGCCTAC 318 |
| QY | 241 atggagctccagcctgacatctctgagattctctgagtccttattctgcaagagatact 300 |
| Db | 319 CTGACCTCAGGAGCCTGACATCTGAGGACACTGCGCTCTATTATTGTAATGAGGGGACT 378 |
| QY | 301 -----acgggtccggttgcttactgggtccaaggacacacagcgtcacgctctcca 351 |
| Db | 379 CCGACTGGCCGCTACTACTTTGACTTGGGGCCAAAGGACACGGGTACCGGTCTCTCTCA 438 |
| QY | 352 ggtggaggcgggttcaggcggagggtggctgctggcgggtggcggatcgagacatcgactcact 411 |
| Db | 439 GGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCAGAAAATGCTCACC 498 |
| QY | 412 cagttccaggaatcatgtctctcatctctccaggggagaaaggtcaacatgacctgcagtgagc 471 |
| Db | 499 CAGTCTCCAGCAATCATGTCTGCATCTCCAGGGAGAGAGTCAACATAACCTGAGTGCC 558 |
| QY | 472 agtcaagtataagttacatgcactgggtaccagcagaagcctgtcacctcccccaaga 531 |
| Db | 559 AGCTCAAGTGTAGTTTACATGCACCTGGTTCAGCAGAGACCCAGGACCTTCTCCCAACTC 618 |
| QY | 532 tggattatgacacatccaactggctctctgagtcctgagtcctgcttcaagtgagtgagg 591 |


```

/strain="BALB/C"
/db_xref="taxon:10090"
/tissue_type="SPLENCYTES"
1. .732
/note="unnamed protein product; protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02927.1"
/db_xref="PID:e306392"
/db_xref="PID:g2300754"
/db_xref="GI:2300754"
/translation="EVQLQSGAEIVKPGASVKLSCKASGYTFTSHMHWVKRQGQ
LEWIFGPNQGRITNVEKFKATLTVDSSTAYMQLSSLTSEDSAVYVCASRYD
YDRFYDFWQGGTIVYSSGGGGGGGGSDIELTQSPAIMASGAPGKVTMTCS
ASSSVYMWYQQKPSPPRLLIYDINSLASGVVPRFSGSGSGTSYLTISRMEADA
ATVYCCQWSSYPMPYIFGGGTKLEIK"
172 a 196 c 209 g 155 t
BASE COUNT
ORIGIN

```

[illegible][illegible]

```

FEATURES             Location/Qualifiers
     1..732
     /organism="Mus musculus"
     /strain="BALE/C"
     /db_xref="taxon:10090"
     /tissue_type="SPLENOCTYES"
     1..732
     /note="unnamed protein product; Protein sequence is in
     conflict with the conceptual translation"
     /codon_start=1
     /protein_id="CAA029333.1"
     /db_xref="PID:e306175"
     /db_xref="PID:g2300766"
     /db_xref="GI:2300766"
     /translation="EVQLQDSAEIYVKPGASVKLSCKASGYTFTSHMHVVKORAGQG
     YDGRFYDSNGRTINERFKSKATLTVDKSSATYQLSSLTSEDSAVYVCASRDYD
     LWDGIFDYGQGTFTVYSGGGSGGGGGGGSDIELTQSPITMSAPEGKVTMTCS
     DSSVSATYMQQGTGSSPRLIYDTSNLASGVFVRFSGGSGTYSLTISRMEADA
     ATYICQWSSYPMPYIFGGGKLEIK"
     175 a      191 c      209 g      157 t
                                *
BASE COUNT
ORIGIN

```

[illegible]

| | | | |
|----|-----|---|-----|
| Qy | 242 | tgaggtccgcagcctgacatctgagattctgcagtctattactgtcaag-agatact | 300 |
| Db | 242 | TGAACTCAGCAGCCTGACATCTGAGGACTCTCGGTCTATTACTGTGCCAGTCGGACT | 301 |
| Qy | 301 | acggtcccg-----tttgcttactgggtccaagggaacacgggtcaocgtctctct | 349 |
| Db | 302 | ATGATTACGACGACGGTACTTTTGACTTACTGGGGCCCAAGGGACCGGTTCACCGTCTCCT | 361 |
| Qy | 350 | caggtggagcgggttcaggcggagggtggtctctgcggtgcgagatcgacatcgagctca | 409 |
| Db | 362 | CAGGTGCGGTGGCTCGGGCGGTGGTGGGTGGCGCGGATCTGACATTGAGCTCA | 421 |
| Qy | 410 | ctcagttccagacaatcatgtctgcattctccaggggagaaagtcacatgacctgcagt | 469 |
| Db | 422 | CCAGTCTCCAACAATCATGCTCGCATCTCCAGGGGAGAAAGTCCACCATGACTCGAGTG | 481 |
| Qy | 470 | gcagctcaagtataaagttacatgcactggtaccagcagaagcctgtcaocttccccaaaa | 529 |
| Db | 482 | ACAGCTCAAGTGTAAAGTTACATGTAAGTGTACACAGCAGAAGACAGGATCTCTCCCAGAC | 541 |
| Qy | 530 | gatggattatgacacatcaaaactggctctcctggagtcctcctgctcagtgagcagt | 589 |
| Db | 542 | TCCTGATTTATGACATCAACACTGGCTTCTGGAGTCCCTGTGCGTTCAGTGGCAGTG | 601 |
| Qy | 590 | ggctcggagaccttattctcacaatcagcagcatggagcctgtagatctgccaatt | 649 |
| Db | 602 | GGTCTGGGACCTCTTACTCTCACAAATCAGCGGAATGGAGGCTGAAGATGTCGCACCT | 661 |
| Qy | 650 | attactggccatcagcggagtagttaccc-----gtcacgtctcgggtcgtggcacacagt | 703 |
| Db | 662 | ATTACTGCCAGCAGTGGAGTAGTTACCCACCCTATGTACAGTTCGGAGGGGGACAAGT | 721 |
| Qy | 704 | tggaaataaaa | 714 |
| Db | 722 | TGGAATAAAA | 732 |

Search completed: November 30, 1999, 14:09:30
Job time: 5097 sec